

10/526256

Rec'd ~~WFO~~ 28 FEB 2005

REC'D 20 OCT 2003

WIPO

PCT

Pl 1076323

# THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office

October 15, 2003

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM  
THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK  
OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT  
APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A  
FILING DATE.

APPLICATION NUMBER: 10/356,736

FILING DATE: January 30, 2003

RELATED PCT APPLICATION NUMBER: PCT/US03/23926



By Authority of the  
COMMISSIONER OF PATENTS AND TRADEMARKS

*W. Montgomery*  
W. MONTGOMERY  
Certifying Officer

**PRIORITY  
DOCUMENT**

SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

**BEST AVAILABLE COPY**

01/30/03  
JCS61 U.S. PRO

92-03-03 10356 6.013003

Albaser

UTILITY PATENT APPLICATION TRANSMITTAL (only for new and continuation-in-part nonprovisional applications under 37 CFR 1.53(b))	Docket No.: P-CE 5480
--	--------------------------

Address to: COMMISSIONER FOR PATENTS  
Box Patent Application  
Washington, D.C. 20231



23601

PATENT TRADEMARK OFFICE

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

"EXPRESS MAIL" MAILING LABEL NUMBER: EV 237 990 811 US

DATE OF DEPOSIT: January 30, 2003

I HEREBY CERTIFY THAT THIS PAPER OR FEE IS BEING  
DEPOSITED WITH THE UNITED STATES POSTAL SERVICE  
"EXPRESS MAIL POST OFFICE TO ADDRESSEE" SERVICE UNDER  
37 C.F.R. 1.10 ON THE DATE INDICATED ABOVE, AND IS  
ADDRESSED TO: COMMISSIONER FOR PATENTS, BOX PATENT APPLICATION,  
WASHINGTON, D.C. 20231.

Christine M. Grace

(TYPED OR PRINTED NAME OF PERSON MAILING PAPER OR FEE)

Christine M. Grace

(SIGNATURE OF PERSON MAILING PAPER OR FEE)

This is a request for filing a  
X new utility patent application under 37 CFR 1.53(b).  
\_\_\_\_ continuation-in-part under CFR 1.53(b)(2) of prior  
application serial no. \_\_\_\_\_, filed \_\_\_\_\_  
(list entire parentage).

Title: MUTATIONS IN NOD2 ARE ASSOCIATED WITH FIBROSTENOSING  
DISEASE IN PATIENTS WITH CROHN'S DISEASE

Inventor(s) (full name of each inventor): Maria T. Abreu, Kent D.  
Taylor, Jerome I. Rotter, Huiying Yang, Kazuhito Sugimura and  
Stephan R. Targan

Enclosed are:

<u>X</u>	Return receipt postcard
<u>X</u>	Patent Application Bibliographic Data Sheet
<u>X</u>	<u>1</u> Page application cover sheet
<u>X</u>	<u>71</u> Pages of specification (includes claims and abstract)
<u>X</u>	<u>15</u> Sheets of drawing(s)
____	Pages of an executed Declaration for Patent Application
____	An executed Power of Attorney for Patent Application by Assignee
<u>X</u>	Paper copy of sequence listing, pages 1 through 19
<u>X</u>	Sequence listing in computer readable form
<u>X</u>	Statement Under 37 CFR 1.821(f)
____	An executed assignment and cover sheet
____	An executed Statement Under 37 CFR 3.73(b)
____	An executed small entity statement
____	Request for Nonpublication and Certification
____	Also enclosed: _____

\_\_\_\_ This application is based on prior foreign application(s)  
No.(s) \_\_\_\_\_, filed in \_\_\_\_\_ on \_\_\_\_\_  
\_\_\_\_\_, respectively, and priority is hereby claimed  
therefrom.

10/356736  
01/30/03  
JCS715 U.S. PRO

Inventors: Abreu et al.  
 Docket No.: P-CE 5480  
 Page 2

X This application is based on, and claims the benefit of, U.S. Provisional Application No. 60/407,391, filed August 30, 2002, and entitled MUTATIONS IN NOD2 ARE ASSOCIATED WITH FIBROSTENOSING DISEASE IN PATIENTS WITH CROHN'S DISEASE, and which is incorporated herein by reference.

— This application is based on, and claims the benefit of, U.S. Provisional Application No. 60/\_\_\_\_\_ (yet to be assigned), filed \_\_\_\_\_, which was converted from U.S. Serial No. \_\_\_\_\_, and entitled \_\_\_\_\_, and which is incorporated herein by reference.

The filing fee has been calculated as shown below:

	Number Filed		Number Extra		Rate		Fee	
					Small Entity	Other Entity	Small Entity	Other Entity
Total Claims	23- 20	=	3	x	\$9	\$18	= \$	\$
Independent Claims	2 - 3	=	0	x	\$42	\$84	= \$	\$
Multiple Dependent Claims Presented: ___ Yes <u>X</u> No					\$140	\$280	\$	\$
					BASIC FEE		\$375	\$750
					TOTAL FEE		\$	\$

— A check in the amount of \$\_\_\_\_\_ to cover the filing fee is enclosed.

X The payment of the filing fee is to be deferred until the Declaration is filed. Do not charge our deposit account.

— The Commissioner is hereby authorized to charge fees under 37 CFR 1.16 and 1.17 which may be required or credit any overpayment to Deposit Account No. \_\_\_\_\_. A duplicate copy of this sheet is enclosed.

Address all future communications to:

Cathryn Campbell  
 CAMPBELL & FLORES LLP  
 4370 La Jolla Village Drive, 7<sup>th</sup> Floor  
 San Diego, California 92122  
 telephone: (858) 535-9001  
 facsimile: (858) 535-8949  
 USPTO CUSTOMER NO. 23601

Inventors: Abreu et al.  
Docket No.: P-CE 5480  
Page 3

Respectfully submitted,

Date: January 30, 2003

Andrea L. Gashler  
Andrea L. Gashler  
Registration No. 41,029  
CAMPBELL & FLORES LLP  
4370 La Jolla Village Dr., 7<sup>th</sup> Fl.  
San Diego, California 92122

PATENT APPLICATION BIBLIOGRAPHIC  
DATA SHEET (3 PAGES)  
Attorney Docket No: P-CE 5480

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

"EXPRESS MAIL" MAILING LABEL NUMBER: EV 237 990 811 US

DATE OF DEPOSIT: January 30, 2003

I HEREBY CERTIFY THAT THIS PAPER OR FEE IS BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE "EXPRESS MAIL POST OFFICE TO ADDRESSEE" SERVICE UNDER 37 CFR 1.10 ON THE DATE INDICATED ABOVE AND IS ADDRESSED TO COMMISSIONER FOR PATENTS, BOX PATENT APPLICATION, WASHINGTON, D.C. 20231.

Christine M Grace  
Printed Name of Person Mailing Paper or Fee

Christine M Grace  
Signature of Person Mailing Paper or Fee

PATENT

Our Docket: P-CE 5480

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: )  
Abreu et al. )

Serial No.: Not Yet Assigned )

Filed: Herewith )

For: MUTATIONS IN NOD2 ARE )  
ASSOCIATED WITH )  
FIBROSTENOSING DISEASE )  
IN PATIENTS WITH CROHN'S )  
DISEASE )

Commissioner for Patents  
Washington, D.C. 20231

Sir:

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

"EXPRESS MAIL" MAILING LABEL NUMBER: EV 237 990 811 US

DATE OF DEPOSIT: January 30, 2003

I HEREBY CERTIFY THAT THIS PAPER OR FEE IS BEING DEPOSITED  
WITH THE UNITED STATES POSTAL SERVICE "EXPRESS MAIL POST  
OFFICE TO ADDRESSEE" SERVICE UNDER 37 CFR 1.10 ON THE DATE  
INDICATED ABOVE AND IS ADDRESSED TO COMMISSIONER FOR  
PATENTS, BOX PATENT APPLICATION, WASHINGTON, D.C. 20231.

Christine M. Grace  
Printed Name of Person Mailing Paper or Fee

Christine M. Grace  
Signature of Person Mailing Paper or Fee

STATEMENT UNDER 37 C.F.R. § 1.821(f)

I hereby state that the content of the paper and  
computer readable copies of the Sequence Listing, submitted in  
accordance with 37 CFR § 1.821(c) and (e), respectively, are the  
same.

Respectfully submitted,

January 30, 2003  
Date

CAMPBELL & FLORES LLP  
4370 La Jolla Village Drive  
7<sup>th</sup> Floor  
San Diego, California 92122  
USPTO CUSTOMER NO. 23601

Andrea L. Gashler  
Andrea L. Gashler  
Registration No. 41,029  
Telephone: (858) 535-9001  
Facsimile: (858) 535-8949

PAPER COPY OF SEQUENCE LISTING,  
PAGES 1 THROUGH 19  
Attorney Docket No: P-CE 5480

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

"EXPRESS MAIL" MAILING LABEL NUMBER: EV 237 990 811 US

DATE OF DEPOSIT: January 30, 2003

I HEREBY CERTIFY THAT THIS PAPER OR FEE IS BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE "EXPRESS MAIL POST OFFICE TO ADDRESSEE" SERVICE UNDER 37 CFR 1.10 ON THE DATE INDICATED ABOVE AND IS ADDRESSED TO COMMISSIONER FOR PATENTS, BOX PATENT APPLICATION, WASHINGTON, D.C. 20231.

Christine McGrace

Printed Name of Person Mailing Paper or Fee

Christine McGrace

Signature of Person Mailing Paper or Fee

INVENTOR INFORMATION

Inventor One Given Name:: Maria T  
Family Name:: Abreu  
Postal Address Line One:: 1309 Marinette Road  
City:: Pacific Palisades  
State or Province:: California  
Country:: United States of America  
Postal or Zip Code:: 90272  
City of Residence:: Pacific Palisades  
State or Province of Residence:: California  
Country of Residence:: United States of America  
Citizenship Country:: United States of America  
Inventor Two Given Name:: Kent D  
Family Name:: Taylor  
Postal Address Line One:: 9302 Halifax Street  
City:: Ventura  
State or Province:: California  
Country:: United States of America  
Postal or Zip Code:: 93004  
City of Residence:: Ventura  
State or Province of Residence:: California  
Country of Residence:: United States of America  
Citizenship Country:: United States of America  
Inventor Three Given Name:: Jerome I  
Family Name:: Rotter  
Postal Address Line One:: 2617 Greenfield Avenue  
City:: Los Angeles  
State or Province:: California  
Country:: United States of America  
Postal or Zip Code:: 90064  
City of Residence:: Los Angeles  
State or Province of Residence:: California  
Country of Residence:: United States of America  
Citizenship Country:: United States of America  
Inventor Four Given Name:: Huiying  
Family Name:: Yang  
Postal Address Line One:: 16409 Holmes Place  
City:: Cerritos  
State or Province:: California  
Country:: United States of America  
Postal or Zip Code:: 90703  
City of Residence:: Cerritos  
State or Province of Residence:: California  
Country of Residence:: United States of America  
Citizenship Country:: United States of America  
Inventor Five Given Name:: Kazuhito  
Family Name:: Sugimura  
Postal Address Line One:: 418 South Genesee Avenue  
City:: Los Angeles



State or Province:: California  
Country:: United States of America  
Postal or Zip Code:: 90036  
City of Residence:: Los Angeles  
State or Province of Residence:: California  
Country of Residence:: United States of America  
Citizenship Country:: Japan  
Inventor Six Given Name:: Stephan R  
Family Name:: Targan  
Postal Address Line One:: 240 22nd Street  
City:: Santa Monica  
State or Province:: California  
Country:: United States of America  
Postal or Zip Code:: 90402  
City of Residence:: Santa Monica  
State or Province of Residence:: California  
Country of Residence:: United States of America  
Citizenship Country:: United States of America

## CORRESPONDENCE INFORMATION

Correspondence Customer Number:: 23601  
Name Line One:: CAMPBELL & FLORES LLP  
Name Line Two:: 4370 La Jolla Village Drive  
Address Line One:: 7th Floor  
City:: San Diego  
State or Province:: California  
Country:: USA  
Postal or Zip Code:: 92122  
Telephone One:: (858) 535-9001  
Fax One:: (858) 535-8949

## APPLICATION INFORMATION

Title Line One:: MUTATIONS IN NOD2 ARE ASSOCIATED WITH FI  
Title Line Two:: BROSTENOSING DISEASE IN PATIENTS WITH CR  
Title Line Three:: OHN'S DISEASE  
Total Drawing Sheets:: 15  
Formal Drawings?: Yes  
Application Type:: Utility  
Docket Number:: P-CE 5480  
License US Govt. Agency:: NIDDK  
Contract or Grant Numbers One:: DK46763  
Secrecy Order in Parent Appl.?: No

## REPRESENTATIVE INFORMATION

Representative Customer Number:: 23601  
Registration Number One:: 31815

Registration Number Two:: 44048  
Registration Number Three:: 39200  
Registration Number Four:: 51566  
Registration Number Five:: 51228  
Registration Number Six:: 41029  
Registration Number Seven:: 47956

CONTINUITY INFORMATION

This application is a:: NON PROV. OF PROVISIONAL  
> Application One:: 60/407,391  
Filing Date:: 08-30-2002

Source:: PrintEFS Version 1.0.1

A P P L I C A T I O N

for

UNITED STATES LETTERS PATENT

on

MUTATIONS IN NOD2 ARE ASSOCIATED WITH FIBROSTENOSING DISEASE  
IN PATIENTS WITH CROHN'S DISEASE

by

Maria T. Abreu

Kent D. Taylor

Jerome I. Rotter

Huiying Yang

Kazuhito Sugimura

and

Stephan R. Targan

Sheets of Drawings: 15

Docket No.: P-CE 5480

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

"EXPRESS MAIL" MAILING LABEL NUMBER: EV 237 990 811 US

DATE OF DEPOSIT: January 30, 2003

I HEREBY CERTIFY THAT THIS PAPER OR FEE IS BEING DEPOSITED WITH THE  
UNITED STATES POSTAL SERVICE "EXPRESS MAIL POST OFFICE TO ADDRESSEE"  
SERVICE UNDER 37 CFR 1.10 ON THE DATE INDICATED ABOVE AND IS  
ADDRESSED TO COMMISSIONER FOR PATENTS, BOX PATENT APPLICATION,  
WASHINGTON, D.C. 20231.

CHRISTINE M GRACE

Printed Name of Person Mailing Paper or Fee

Christine M Grace

Signature of Person Mailing Paper or Fee

Attorneys

CAMPBELL & FLORES LLP

4370 La Jolla Village Drive, Suite 700

San Diego, California 92122

USPTO CUSTOMER NO. 23601

MUTATIONS IN NOD2 ARE ASSOCIATED WITH FIBROSTENOSING DISEASE  
IN PATIENTS WITH CROHN'S DISEASE

RELATED APPLICATIONS

5

This application claims the benefit of priority of provisional application serial number 60/407,391, filed August 30, 2002, and which is incorporated herein by reference.

10

ACKNOWLEDGMENT

This work was supported by grant DK46763 and DK54967 awarded by NIDDK. The United States government has certain rights in this invention.

15

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

20

The invention relates generally to the fields of genetics and autoimmune disease and, more specifically, to mutations linked to the NOD2/CARD15 gene and genetic methods for diagnosing clinical subtypes of Crohn's disease.

25

BACKGROUND INFORMATION

Inflammatory bowel disease (IBD) is the collective term used to describe two gastrointestinal disorders of unknown etiology: Crohn's disease (CD) and ulcerative colitis (UC). The course and prognosis of IBD, which occurs world-wide and is reported to afflict as many as two million

30

people, varies widely. Onset of IBD is predominantly in young adulthood with diarrhea, abdominal pain, and fever the three most common presenting symptoms. The diarrhea may range from mild to severe, and anemia and weight loss are additional common signs of IBD. Of all patients with IBD, ten percent to fifteen percent will require surgery over a ten year period. In addition, patients with IBD are at increased risk for the development of intestinal cancer. Reports of an increasing occurrence of psychological problems, including anxiety and depression, are perhaps not surprising symptoms of what is often a debilitating disease that strikes people in the prime of life.

Crohn's disease is a classification representing a number of heterogeneous disease subtypes that affect the gastrointestinal tract and produce similar symptoms. Both environmental and genetic factors likely contribute to the etiology of such disease subtypes. Patients with Crohn's disease can be classified, for example, into subtypes based on the presence of fibrostenosing disease, internal-perforating disease, perianal fistulizing disease or ulcerative colitis-like disease according to previously described criteria. The fibrostenosing disease subtype is characterized by documented persistent intestinal obstruction or intestinal resection for intestinal obstruction. The extensive and often protracted clinical testing required to diagnose Crohn's disease and disease subtypes may delay optimal treatment and involves invasive procedures such as endoscopy.

Identification of genetic markers which are closely associated with a clinical subtype of Crohn's disease would provide the basis for novel genetic tests and

eliminate or reduce the need for the battery of laboratory, radiological, and endoscopic evaluations typically required to determine disease subtype. The availability of methods for diagnosing clinical subtypes of Crohn's disease would represent a major clinical advance that would aid in the therapeutic management of Crohn's disease and would further lay the groundwork for the design of treatment modalities which are specific to a particular disease subtype. Such methods can reduce costs associated with treatment of unresponsive disease subtypes and eliminate the disappointment of those needlessly undergoing ineffective therapy. In particular, a reliable genetic test for the fibrostenosing subtype of Crohn's disease would be highly prized as a non-invasive method for the early diagnosis of this disease subtype and would also be useful for predicting susceptibility to the fibrostenosing subtype of Crohn's disease in asymptomatic individuals, making prophylactic therapy possible. The present invention satisfies this need and provides related advantages as well.

#### SUMMARY OF THE INVENTION

The present invention provides a method of diagnosing or predicting susceptibility to a clinical subtype of Crohn's disease characterized by fibrostenosing disease by determining the presence or absence in an individual of a fibrostenosis-predisposing allele linked to a NOD2/CARD15 locus, where the presence of the fibrostenosis-predisposing allele is diagnostic of or predictive of susceptibility to the clinical subtype of Crohn's disease characterized by fibrostenosing disease. In a method of the invention, the clinical subtype of Crohn's

disease can be, for example, characterized by fibrostenosing disease independent of small bowel involvement.

5 The invention also provides a method of optimizing therapy in an individual by determining the presence or absence in the individual of a fibrostenosis-predisposing allele linked to a NOD2/CARD15 locus, diagnosing individuals in which the fibrostenosis-predisposing allele is present as having a fibrostenosing subtype of Crohn's disease, and  
10 treating the individual having a fibrostenosing subtype of Crohn's disease based on the diagnosis.

#### BRIEF DESCRIPTION OF THE DRAWINGS

15 Figure 1 shows the NOD2/CARD15 locus intron and exon structure with the location of SNP 8, SNP 12, SNP 13, and JW1 as well as other markers.

20 Figure 2 shows the frequency of NOD2 variant carriers having at least one of the three NOD2/CARD15 rare variant alleles (a "2" allele at SNP 8, SNP 12 or SNP 13) in Cohort 1 ("CD1"), Cohort 2 ("CD2") or a combination of Cohort 1 and Cohort 2 ("Combined CD"). The striped bars indicate Crohn's disease patients with fibrostenosing  
25 disease and the solid bars indicate Crohn's disease patients who did not have the fibrostenosing subtype of disease.

30 Figure 3 shows the frequency of fibrostenosing complications in patients relative to the number of NOD2/CARD15 rare variant alleles at SNP 8, SNP 12 or SNP 13. Based on genotyping at SNP 8, SNP 12 and SNP 13, patients were described as carrying 0, 1, or 2 rare variant alleles

(x axis, number of mutant NOD2 alleles). The left y axis shows the frequency of fibrostenosing complications as filled circles. The right y axis shows the odds ratio as a filled diamond with 95% confidence intervals in parentheses. The \* indicates a p value = 0.008 and \*\* indicates a p value of 0.004 compared with 0 alleles.

Figure 4 shows a comparison of NOD2/CARD15 variant allelic frequencies in patients with fibrostenosing disease compared with perforating disease. Patients were separated by the presence of fibrostenosing disease with perforating complications ("Fib + perf") or fibrostenosing disease without perforating complications ("Fib only") compared with patients with perforating complications and without evidence of fibrostenosis ("Perf only").

Figure 5 shows the nucleotide sequence of NOD2/CARD15 surrounding SNP 8. The top strand is labeled as SEQ ID NO:1 and the bottom strand is labeled as SEQ ID NO:2. Nucleotide sequences which can be used as primers for PCR amplification are indicated. In addition, the position of a nucleotide sequence which can be used as a probe in an allelic discrimination assay is indicated, in this figure, by a box and lower-case letters. The underlined nucleotide indicates the position of the polymorphic site.

Figure 6 shows the nucleotide sequence of NOD2/CARD15 surrounding SNP 12. The top strand is labeled as SEQ ID NO:3 and the bottom strand is labeled as SEQ ID NO:4. Nucleotide sequences which can be used as primers for PCR amplification are indicated. In addition, the position of a nucleotide sequence which can be used as a probe in an allelic discrimination assay is indicated, in this figure,



by a box and lower case letters. The underlined nucleotide indicates the position of the polymorphic site.

Figure 7 shows the nucleotide sequence of  
5 NOD2/CARD15 surrounding SNP 13. The top strand is labeled  
as SEQ ID NO:5 and the bottom strand is labeled as SEQ ID  
NO:6. Nucleotide sequences which can be used as primers for  
PCR amplification are indicated. In addition, the position  
of a nucleotide sequence which can be used as a probe in an  
10 allelic discrimination assay is indicated, in this figure,  
by a box and lower case letters. The underlined nucleotide  
indicates the position of the polymorphic site.

Figure 8 shows the nucleotide sequence of  
15 NOD2/CARD15 surrounding SNP 5. The top strand is labeled as  
SEQ ID NO:7 and the bottom strand is labeled as SEQ ID NO:8.  
Nucleotide sequences which can be used as primers for PCR  
amplification are indicated. In addition, the position of a  
nucleotide sequence which can be used as a probe in an  
20 allelic discrimination assay is indicated, in this figure,  
by a box and lower case letters. The underlined nucleotide  
indicates the position of the polymorphic site.

Figure 9 shows the nucleotide sequence of  
25 NOD2/CARD15 surrounding the JW1 variant sequence. The top  
strand is labeled as SEQ ID NO:9 and the bottom strand is  
labeled as SEQ ID NO:10. Nucleotide sequences which can be  
used as primers for PCR amplification are indicated. In  
addition, the position of a nucleotide sequence which can be  
30 used as a probe in an allelic discrimination assay is  
indicated, in this figure, by a box and lower case letters.  
The underlined nucleotide indicates the position of the  
polymorphic site.

Figure 10 shows the nucleotide sequence of the 5' untranslated region of NOD2/CARD15 in 12 individuals (SEQ ID NOS:12-23) compared to the wild-type NOD2/CARD15 sequence (SEQ ID NO:11). Areas of sequence identity are shaded. The location of two polymorphic sites, JW18 and JW17, are indicated.

Figure 11 shows the nucleotide sequence of the 3' untranslated region of NOD2/CARD15 in 12 individuals. Areas of sequence identity are shaded. The location of two polymorphic sites, JW15 and JW16, are indicated.

Figure 11 A shows the nucleotide sequence of the 3' untranslated region of NOD2/CARD15 in 12 individuals (SEQ ID NOS:25-36) compared to the wild-type NOD2/CARD15 sequence (SEQ ID NO:24) and the location of JW16. Figure 11 B shows the nucleotide sequence of the 3' untranslated region of NOD2/CARD15 in 12 individuals (SEQ ID NOS:56-67) compared to the wild-type NOD2/CARD15 sequence (SEQ ID NO:55) and the location of JW15.

20

#### DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed to the exciting discovery of disease-predisposing alleles that are closely associated with the fibrostenosing disease subtype of Crohn's disease. These fibrostenosis-predisposing alleles are linked to a NOD2/CARD15 locus as described further below and can be used to diagnose or predict susceptibility to the fibrostenosing disease subtype of Crohn's disease.

30

As disclosed herein, genotyping and other clinical characterization approaches were used to identify a strong

association between disease-predisposing alleles and the fibrostenosing disease subtype of Crohn's disease. In particular, two cohorts of Crohn's disease patients were assembled and clinically characterized (see Example I and Table 1). These patients were also genotyped for three single nucleotide polymorphisms (SNPs) in the NOD2/CARD15 gene, SNP 8, SNP 12, and SNP 13, which are polymorphic markers associated with Crohn's disease. As disclosed herein, univariate analysis indicated that a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus was significantly associated with fibrostenosing disease in Cohort 1 ( $p=0.049$ , see Table 5). In addition, a positive association at a less stringent significance level was also observed with small bowel involvement and younger age of onset, and a negative association was observed with ulcerative colitis-like disease in this cohort. With respect to serologic markers, patients with the "2" allele at SNP 13 were more likely to express anti-*Saccharomyces cerevisiae* antibodies (ASCA) ( $p=0.053$ ).

The results obtained with Cohort 1 were further tested using Cohort 2 as disclosed herein in Example IV. As with Cohort 1, Cohort 2 demonstrated a significant association between a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus and fibrostenosing disease ( $p=0.002$ , see Table 6). Furthermore, the significance between a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus and fibrostenosing disease increased when the two cohorts were analyzed together ( $p=0.001$ , see Figure 2).

As further disclosed herein in Example V, a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus

was associated with fibrostenosing disease in both Jewish and non-Jewish individuals. Approximately 46% of Crohn's disease patients with fibrostenosing disease (Jewish individuals 52% vs. non-Jewish individuals 42%) had at least one of these rare variant alleles compared with only 23% (Jewish individuals 21.6% vs. non-Jewish individuals 25%) of Crohn's disease patients without fibrostenosing disease (Odds ratio, 2.8; 95% Confidence interval, 1.56-5.18). Of the three rare variant alleles, the "2" allele at SNP 13, which is a frameshift mutation denoted "3020insC," demonstrated the greatest association with fibrostenosing disease (47% vs. 17%,  $p=0.006$  for cohorts combined). These results indicate that fibrostenosis-predisposing alleles can be linked to the NOD2/CARD15 locus.

As further disclosed herein in Example VI, patients who were carriers of two fibrostenosis-predisposing alleles in NOD2/CARD15 were significantly more likely to have fibrostenosing disease as compared with patients who were not carriers of NOD2/CARD15 mutations at SNP 8, SNP 12 or SNP 13 (85% vs. 43%; odds ratio 7.4; 95% confidence interval 1.9-28.9,  $p=0.004$ ). See Figure 3. Patients who were carriers of a single NOD2/CARD15 fibrostenosis-predisposing allele were also significantly more likely to have fibrostenosing disease when compared with patients who were not carriers of any of the three NOD2/CARD15 fibrostenosis-predisposing alleles assayed (64% vs. 43%; odds ratio 2.37; 95% confidence interval 1.26-4.47;  $p=0.008$ ). These results confirm that patients who have a fibrostenosis-predisposing allele linked to a NOD2/CARD15 locus can have the fibrostenosing subtype of Crohn's disease and further indicate that Crohn's disease patients with multiple fibrostenosis-predisposing alleles (homozygous

mutations or compound heterozygous mutations) linked to NOD2/CARD15 have an increased risk of fibrostenosing disease as compared to individuals carrying a single fibrostenosis-predisposing allele.

5

Fibrostenosing and perforating disease can occur together in the same patient. Patients with fibrostenosing disease can be characterized, for example, as i) having only fibrostenosing disease or ii) having both fibrostenosing and perforating disease. As disclosed herein in Example VII, the percentage of patients having only fibrostenosing disease that carried a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus was 48.3%, which was similar to that seen in patients with both fibrostenosing and perforating complications (46.0%;  $p=0.8$ ). As seen in Figure 4, when patients with fibrostenosing disease were compared with those patients described as having perforating disease only (perianal or internal), the frequency of the "2" allele at SNP 8, SNP 12 or SNP 13 of the NOD2/CARD15 locus in patients with fibrostenosing disease (with or without perforating complications) was significantly greater than that seen in patients with only perforating complications (46.6% versus 18.6%;  $p=0.002$ ).

As further disclosed herein in Example VIII, multivariant analysis was performed to investigate the association of a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus with clinical phenotypes. For multivariant analysis, all variables with at least borderline significance ( $p < 0.1$ ) in either cohort were tested simultaneously for their association with a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus using logistic regression. As shown in Table 7, the

clinical phenotype of fibrostenosing disease was significantly associated ( $p < 0.05$ ) with these rare alleles at the NOD2/CARD15 locus (odds ratio 2.8; 95% confidence interval, 1.3-6.0). These results confirm that

5 fibrostenosing disease is independently associated with a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus.

Because fibrostenosing disease is more likely to

10 occur in patients with small-bowel involvement, patients were stratified on the basis of small-bowel involvement to analyze whether the association between fibrostenosing disease and NOD2/CARD15 variant alleles was a primary association (see Example IX). Among patients with

15 small-bowel involvement, 26.4% of patients who did not have fibrostenosing disease ( $n=53$ ) had a "2" allele at SNP 8, SNP 12, or SNP 13, whereas 46.1% of patients who had fibrostenosing disease ( $n=102$ ) had a "2" allele at SNP 8, SNP 12, or SNP 13 ( $p=0.017$ ). A similar trend was observed

20 among patients without small-bowel involvement ( $p=0.05$ ), and the combined analysis conditioning on small-bowel involvement yielded a significance level of 0.009. After controlling for fibrostenosing disease, small-bowel involvement was not associated with a "2" allele at SNP 8,

25 SNP 12, or SNP 13 of the NOD2/CARD15 locus ( $p=0.63$ ). This result agrees with the results from logistic regression analysis (see Example VIII) and indicates that the association between fibrostenosing disease and a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus is

30 independent of small-bowel involvement. These results further indicate that the observed small-bowel association with a "2" allele at SNP 8, SNP 12, or SNP 13 of the

NOD2/CARD15 locus is secondary to the presence of fibrostenosing disease.

Based on these discoveries, the present invention provides a method of diagnosing or predicting susceptibility to a clinical subtype of Crohn's disease characterized by fibrostenosing disease by determining the presence or absence in an individual of a fibrostenosis-predisposing allele linked to a NOD2/CARD15 locus, where the presence of the fibrostenosis-predisposing allele is diagnostic of or predictive of susceptibility to the clinical subtype of Crohn's disease characterized by fibrostenosing disease. The methods of the invention can be advantageous in that they are noninvasive and can be conveniently practiced, for example, with a blood sample from an individual. The methods of the invention can be used to quickly, easily and reliably diagnose or predict susceptibility to a clinical subtype of Crohn's disease characterized by fibrostenosing disease as described further herein below.

In one embodiment, a method of the invention is practiced with a fibrostenosis-predisposing allele located within the NOD2/CARD15 locus. In another embodiment, NF-kappa B activation by a NOD2/CARD15 polypeptide encoded by the fibrostenosis-predisposing allele is reduced as compared to NF-kappa B activation by a wild-type NOD2/CARD15 polypeptide. In a further embodiment, a method of the invention is practiced with a fibrostenosis-predisposing allele located in a coding region of the NOD2/CARD15 locus, for example, within a region encoding residues 744 to 1020 of NOD2/CARD15. In still a further embodiment, a method of the invention is practiced with a fibrostenosis-predisposing allele which is a "2" allele at SNP 8, SNP 12, or SNP 13.

In yet a further embodiment, a method of the invention is practiced with a fibrostenosis-predisposing allele which is a "2" allele at SNP 13. In another embodiment, a method of the invention is practiced with a fibrostenosis-predisposing allele which is located in a non-coding region of the NOD2/CARD15 locus. Such an allele can be, without limitation, a JW1, JW15, or JW16 variant allele. In another embodiment, a method of the invention is practiced with a fibrostenosis-predisposing allele which is located in a promoter region of the NOD2/CARD15 locus. Useful fibrostenosis-predisposing alleles in the promoter region of NOD2/CARD15 include, but are not limited to, JW17 and JW18 variant alleles.

The present invention relates to genetic markers which localize to the IBD1 locus on chromosome 16. Utilizing genome wide scan linkage strategies, the IBD1 locus was mapped to the proximal region of the long arm of chromosome 16 (16q12) in the Caucasian population (Hugot et al., Nature 379:821-823 (1996)). This finding has been replicated in many studies, including an international collaborative study reporting a high multipoint linkage score (MLS) for a complex disease (MLS = 5.7 at marker D16S411 in 16q12). See Cho et al., Inflamm. Bowel Dis. 3:186-190 (1997), Akolkar et al., Am. J. Gastroenterol. 96:1127-1132 (2001), Ohmen et al., Hum. Mol. Genet. 5:1679-1683 (1996), Parkes et al., Lancet 348:1588 (1996), Cavanaugh et al., Ann. Hum. Genet. (1998), Brant et al., Gastroenterology 115:1056-1061 (1998), Curran et al., Gastroenterology 115:1066-1071 (1998), Hampe et al., Am. J. Hum. Genet. 64:808-816 (1999), and Annese et al., Eur. J. Hum. Genet. 7:567-573 (1999). NOD2/CARD15 within the IBD1 locus was simultaneously identified by a positional-cloning



strategy (Hugot et al., Nature 411:599-603 (2001)) and a positional candidate gene strategy (Ogura et al., Nature 411:603-606 (2001), Hampe et al., Lancet 357:1925-1928 (2001)). The encoded NOD2/CARD15 protein contains amino-terminal caspase recruitment domains (CARDs), which can activate NF-kappa B (NF-kB), and several carboxy-terminal leucine-rich repeat domains (Ogura et al, J. Biol. Chem. 276:4812-4818 (2001)). Figure 1 shows an illustration of the NOD2/CARD15 locus.

The sequence of the human NOD2/CARD15 gene can be found in GenBank as accession number NM\_022162, which is incorporated by reference herein. In addition, the complete sequence of human chromosome 16 clone RP11-327F22, which includes NOD2/CARD15, can be found in GenBank as accession number AC007728, which is incorporated by reference herein. Furthermore, the sequence of NOD2/CARD15 from other species can be found in the GenBank database.

Variations at three single nucleotide polymorphisms in the coding region of NOD2/CARD15 have been previously described. These three SNPs, designated SNP 8, SNP 12 and SNP 13, are located in the carboxy-terminal region of the NOD2/CARD15 gene (Hugot et al., *supra*, 2001).

The invention relies, in part, on determining the presence or absence in an individual of a fibrostenosing-predisposing allele linked to a NOD2/CARD15 locus. The term "fibrostenosis-predisposing allele," as used herein, means a stably heritable molecular variation that is linked to the NOD2/CARD15 locus and that tends to be inherited together with the clinical subtype of Crohn's disease characterized

by fibrostenosing disease more often than would be expected according to traditional Mendelian genetics. A fibrostenosis-predisposing allele useful in the invention can be, without limitation, a single nucleotide polymorphism (SNP), microsatellite (ms), variable number tandem repeat (VNTR) polymorphism, or a substitution, insertion or deletion of one or more nucleotides. One skilled in the art understands that a fibrostenosis-predisposing allele also can be a molecular variation such as abnormal methylation or other modification that does not produce a difference in the primary nucleotide sequence of the fibrostenosis-predisposing allele as compared to the wild type allele.

The term "linked to," as used herein in reference to a fibrostenosis-predisposing allele and a NOD2/CARD15 locus, means that the fibrostenosis-predisposing allele and the NOD2/CARD15 locus are inherited together more often than would be expected according to traditional Mendelian genetics. It is understood that an allele and locus are linked when there is less than 50% recombination between the two sites. A fibrostenosis-predisposing allele is generally separated from a NOD2/CARD15 locus by at most 50 centiMorgan (cM). As non-limiting examples, a fibrostenosis-predisposing allele can be within 50 centiMorgan (cM), 40 cM, 30 cM, 20 cM, 10 cM, 5 cM, or 1 cM of the NOD2/CARD15 locus. The distance between a linked fibrostenosis-predisposing allele and a NOD2/CARD15 locus can also be, for example, 50,000,000 base pairs (bps), 40,000,000 bps, 30,000,000 bps, 20,000,000 bps, 10,000,000 bps, 5,000,000 bps, 1,000,000 bps, 500,000 bps, 100,000 bps, 50,000 bps, 10,000 bps, 1,000 bps or 100 bps.

The methods of the invention are useful for diagnosing or predicting susceptibility to a clinical subtype of Crohn's disease (regional enteritis), which is a disease of chronic inflammation that can involve any part of the gastrointestinal tract. Commonly the distal portion of the small intestine (ileum) and cecum are affected in Crohn's disease. In other cases, the disease is confined to the small intestine, colon or anorectal region. Crohn's disease occasionally involves the duodenum and stomach, and more rarely the esophagus and oral cavity.

The variable clinical manifestations of Crohn's disease are, in part, a result of the varying anatomic localization of the disease. The most frequent symptoms of Crohn's disease are abdominal pain, diarrhea and recurrent fever. Crohn's disease is commonly associated with intestinal obstruction or fistula, which is an abnormal passage, for example, between diseased loops of bowel. Crohn's disease also can include extra-intestinal complications such as inflammation of the eye, joints and skin; liver disease; kidney stones or amyloidosis; and is associated with an increased risk of intestinal cancer.

Several features are characteristic of the pathology of Crohn's disease. The inflammation associated with Crohn's disease, known as transmural inflammation, involves all layers of the bowel wall. Thickening and edema, for example, typically appear throughout the bowel wall, with fibrosis also present in long-standing disease. The inflammation characteristic of Crohn's disease also is discontinuous, with segments of inflamed tissue, known as "skip lesions," separated by apparently normal intestine. Furthermore, linear ulcerations, edema, and inflammation of

the intervening tissue lead to a "cobblestone" appearance of the intestinal mucosa, which is distinctive of Crohn's disease.

5           A hallmark of Crohn's disease is the presence of discrete aggregations of inflammatory cells, known as granulomas, which are generally found in the submucosa. About half of Crohn's disease cases display the typical discrete granulomas, while others show a diffuse  
10           granulomatous reaction or nonspecific transmural inflammation. As a result, the presence of discrete granulomas is indicative of Crohn's disease, although the absence granulomas also is consistent with the disease. Thus, transmural or discontinuous inflammation, rather than  
15           the presence of granulomas, is a preferred diagnostic indicator of Crohn's disease (Rubin and Farber, Pathology (Second Edition) Philadelphia: J.B. Lippincott Company (1994)).

20           Crohn's disease is a classification representing a number of heterogeneous disease subtypes that affect the gastrointestinal tract and may produce similar symptoms. As non-limiting examples, patients with Crohn's disease can be characterized as having subtypes characterized by  
25           fibrostenosing disease, internal-perforating disease, perianal fistulizing disease or ulcerative colitis (UC)-like disease based on previously described criteria (Gasche et al., Inflammatory Bowel Diseases 6:8-15 (2000); Vasiliasukas et al., Gut 47:487-496 (2000); Vasiliasukas et al.,  
30           Gastroenterology 110:1810-1819 (1996); and Greenstein et al., Gut 29:588-592 (1988)).

According to well-established criteria fibrostenosing disease is defined by documented persistent intestinal obstruction or an intestinal resection for an intestinal obstruction. Radiographic, endoscopic, surgical or histopathological documentation can be used to confirm a diagnosis of the fibrostenosing subtype of Crohn's disease. The fibrostenosing subtype of Crohn's disease can be accompanied by other symptoms such as perforations, abscesses or fistulae. In addition, the fibrostenosing subtype of Crohn's disease can be characterized by persistent symptoms of intestinal blockage such as nausea, vomiting, abdominal distention and inability to eat solid food. Intestinal X-rays of patients with the fibrostenosing subtype of Crohn's disease can show, for example, distention of the bowel before the point of blockage.

Additional subtypes of Crohn's disease can be identified using defined clinical criteria. For example, internal perforating disease can be defined as current or previous evidence of entero-enteric or entero-vesicular fistulae, intraabdominal abscesses, or small bowel perforation. Perianal perforating disease can be defined by current or previous evidence of either perianal fistulae or abscesses or rectovaginal fistula. UC-like disease can be defined by current or previous evidence of left-sided colonic involvement, symptoms of bleeding or urgency, and crypt abscesses on colonic biopsies as previously described. Disease location can be classified as small bowel, colon, or both based on one or more endoscopic, radiologic or pathologic studies.

The fibrostenosing subtype of Crohn's disease can occur in patients having disease with small-bowel involvement. As disclosed herein, after controlling for fibrostenosing disease, small-bowel involvement was not associated with a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus ( $p=0.63$ ; see Example IX). This result agrees with the results from logistic regression analysis disclosed in Example VIII and indicates that the association between fibrostenosing disease and a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus is independent of small-bowel involvement. Based on these findings, the present invention provides a method of diagnosing or predicting susceptibility to a clinical subtype of Crohn's disease characterized by fibrostenosing disease independent of small bowel involvement by determining the presence or absence in an individual of a fibrostenosis-predisposing allele linked to a NOD2/CARD15 locus, where the presence of the fibrostenosis-predisposing allele is diagnostic of or predictive of susceptibility to the clinical subtype of Crohn's disease characterized by fibrostenosing disease independent of small bowel involvement.

The diagnostic methods of the invention are practiced in an individual. As used herein, the term "individual" means an animal, such as a human or other mammal, capable of having the fibrostenosing subtype of Crohn's disease. An individual can have one or more symptoms of Crohn's disease or the fibrostenosing subtype of Crohn's disease or can be asymptomatic. The methods of the invention can be useful, for example, for diagnosing the fibrostenosing subtype of Crohn's disease in an individual with one or more symptoms, or for predicting susceptibility to the fibrostenosing subtype of Crohn's disease in an

asymptomatic individual such as an individual at increased risk for having the fibrostenosing subtype of Crohn's disease. In one embodiment, the methods of the invention are used to determine the presence or absence of the fibrostenosing subtype of Crohn's disease in an individual known to have Crohn's disease.

The methods of the invention can be useful, for example, to diagnose or predict susceptibility to the fibrostenosing subtype of Crohn's disease in an Ashkenazi Jewish individual. Crohn's disease is significantly more common (2 to 8 fold higher) in Ashkenazi Jews than in non-Jewish Caucasians (Brant et al., Gastroenterol. 115:1056-1061 (1998)). Furthermore, among persons of Jewish ethnicity, American or European Ashkenazi Jews have a 2 to 4 fold increased risk of having this inflammatory bowel disease compared with Sephardic or Oriental Jews (Yang and Rotter in Kirschner and Shorter (Eds.), Inflammatory Bowel Disease Baltimore: Williams and Wilkins, p. 301-331 (1995); Rozen et al., Gastroenterol. 76:25-30 (1979)). The empiric risk of Crohn's disease for a first degree relative of a proband with Crohn's disease is 7.8% for Jews compared with 5.2% for non-Jews (p=0.005; Yang et al., Gut 34:517-524 (1993)). Thus, the Jewish population and especially the Ashkenazi Jewish population represents a group at increased risk for Crohn's and autoimmune diseases of related etiology.

The methods of the invention rely on fibrostenosis-predisposing alleles linked to the NOD2/CARD15 locus. NOD2/CARD15 has structural homology with both apoptosis regulators Apaf-1/CED-4 and a class of plant disease resistant gene products (Ogura et al., J. Biol.

Chem, 276:4812-4818 (2001)). Similar to plant disease  
resistant gene products, NOD2/CARD15 has an amino-terminal  
effector domain, a nucleotide-binding domain and leucine  
rich repeats (LRRs). Wild-type NOD2/CARD15 activates  
5 nuclear factor NF-kappa B, making it responsive to bacterial  
lipopolysaccharides (LPS; Ogura et al., J. Biol. Chem.,  
276:4812-4818 (2001); Inohara et al., J. Biol. Chem.  
276:2551-2554 (2001). NOD2/CARD15 can function as an  
intercellular receptor for LPS, with the leucine rich  
10 repeats required for responsiveness. Like NOD2/CARD15,  
cytosolic plant disease resistant polypeptides have  
carboxy-terminal leucine rich repeats that are important for  
the recognition of pathogen components and induction of  
pathogen-specific responses (Parniske et al., Cell  
15 91:821-832 (1997); Ellis et al., Plant Cell 11:495-506  
(1999); Dixon et al., Proc. Natl. Acad. Sci. USA 97: 8807-  
8814 (2000)).

In one embodiment, the fibrostenosis-predisposing  
20 allele is located within the NOD2/CARD15 locus, a schematic  
of which is shown in Figure 1. The NOD2/CARD15 locus  
includes coding regions of the NOD2/CARD15 gene as well as  
non-coding regions such as introns and 5' and 3'  
untranslated regions. One skilled in the art understands  
25 that the NOD2/CARD15 locus can include, for example,  
promoter regions 5' of the gene, enhancer regions 5' or 3'  
of the gene or in intronic sequences, and mRNA stability  
regions 3' of the gene.

30 In another embodiment, the  
fibrostenosis-predisposing allele is located in a coding  
region of the NOD2/CARD15 locus, for example within a region  
encoding residues 744 to 1020 of NOD2/CARD15. Residues 744



to 1020 of the NOD2/CARD15 polypeptide contain several leucine-rich repeats in the carboxy-terminal portion of the NOD2/CARD15 polypeptide. Fibrostenosis-predisposing alleles located in a region encoding residues 744 to 1020 of NOD2/CARD15 include, without limitation, SNP 12 and SNP 13. A fibrostenosis-predisposing allele useful in the invention also can be an allele which encodes a NOD2/CARD15 polypeptide with reduced NF-kappa B activation as compared to NF-kappa B activation by a wild-type NOD2/CARD15 polypeptide. As an example, a rare variant allele at SNP 13 results in a truncated NOD2/CARD15 polypeptide which has reduced ability to induce NF-kappa B in response to LPS stimulation (Ogura et al., Nature 411:603-606 (2001)).

A fibrostenosis-predisposing allele useful in the invention can be, for example, a "2" allele at SNP 8, SNP 12 or SNP 13. SNP 8, SNP 12, and SNP 13 are located within the coding region of NOD2/CARD15 as shown in Figure 1. In one embodiment, a method of the invention is practiced with a fibrostenosis-predisposing allele which is the SNP 8 "2" allele. As used herein, the term "SNP 8" means a single nucleotide polymorphism within exon 4 in the NOD2/CARD15 gene, which occurs within a triplet encoding amino acid 702 of the NOD2/CARD15 protein. The "1" allele, in which cytosine (c) resides at position 138,991 of the AC007728 sequence, is the common or "wild-type" SNP 8 allele and occurs within a triplet encoding an arginine at amino acid 702. The "2" allele of SNP 8, in which thymine (t) resides at position 138,991 of the AC007728 sequence, is a rare variant that results in an arginine (R) to tryptophan (W) substitution at amino acid 702 of the NOD2/CARD15 protein. Accordingly, the rare "2" allele at SNP 8 is denoted "R702W" or "702W" and can also be denoted "R675W"

based on the earlier numbering system of Hugot et al.,  
*supra*, 2001. The NCBI SNP ID number for SNP 8 is rs2066844,  
which is incorporated herein by reference. As disclosed  
herein and described further below, the presence of allele  
5 "1" or "2" at SNP 8, or another SNP described below, can be  
conveniently detected, for example, by allelic  
discrimination assays or sequence analysis.

A method of the invention also can be practiced  
10 with a fibrostenosis-predisposing allele which is the SNP 12  
"2" allele. As used herein, the term "SNP 12" means a  
single nucleotide polymorphism within exon 8 in the  
NOD2/CARD15 gene, which occurs within a triplet encoding  
amino acid 908 of the NOD2/CARD15 protein. The "1" allele,  
15 in which guanine (g) resides at position 128,377 of the  
AC007728 sequence, is the common or "wild-type" SNP 12  
allele and occurs within a triplet encoding glycine at amino  
acid 908. The "2" allele of SNP 12, in which cytosine (c)  
resides at position 128,377 of the AC007728 sequence, is a  
20 rare variant that results in a glycine (G) to arginine (R)  
substitution at amino acid 908 of the NOD2/CARD15 protein.  
This rare "2" allele at SNP 12 is denoted "G908R" or "908R"  
and can also be denoted "G881R" based on the earlier  
numbering system of Hugot et al., *supra*, 2001. SNP 12 is  
25 located within the leucine rich repeat region of the  
NOD2/CARD15 gene. The NCBI SNP ID number for SNP 12 is  
rs2066845, which is incorporated herein by reference.

A method of the invention also can be practiced  
30 with a fibrostenosis-predisposing allele which is the SNP 13  
"2" allele. This allele is an insertion of a single  
nucleotide that results in a frame shift in the tenth  
leucine-rich repeat of the NOD2/CARD15 protein and is

followed by a premature stop codon. The resulting truncation of the NOD2/CARD15 protein appears to prevent activation of NF- $\kappa$ B in response to bacterial lipopolysaccharides (Ogura et al., *supra*, 2001). As used herein, the term "SNP 13" means a single nucleotide polymorphism within exon 11 in the NOD2/CARD15 gene, which occurs in a triplet encoding amino acid 1007 of the NOD2/CARD15 protein. The "2" allele of SNP 13, in which a cytosine has been added at position 121,139 of the AC007728 sequence, is a rare variant resulting in a frame shift mutation at amino acid 1007. Accordingly, the rare "2" allele at SNP 13 is denoted "1007fs" and can also be denoted "3020insC," or "980fs" based on the earlier numbering system of Hugot et al., *supra*, 2001. The NCBI SNP ID number for SNP 13 is rs2066847, which is incorporated herein by reference.

One skilled in the art recognizes that a particular polymorphic allele can be conveniently defined, for example, in comparison to a Centre d'Etude du Polymorphisme Humain (CEPH) reference individual such as the individual designated 1347-02 (Dib et al., Nature 380:152-154 (1996)), using commercially available reference DNA obtained, for example, from PE Biosystems (Foster City, CA). In addition, specific information on SNPs can be obtained from the dbSNP of the National Center for Biotechnology Information (NCBI).

A fibrostenosis-predisposing allele also can be located in a non-coding region of the NOD2/CARD15 locus. Non-coding regions include, for example, intron sequences as well as 5' and 3' untranslated sequences. Examples of a fibrostenosis-predisposing allele in an intron sequence of

the NOD2/CARD15 gene include, but are not limited to, the JW1 variant allele, which is described below. Examples of fibrostenosis-predisposing alleles located in the 3' untranslated region of the NOD2/CARD15 gene include, without limitation, JW15 and JW16 variant alleles, which are described below. Examples of fibrostenosis-predisposing alleles located in the 5' untranslated region of the NOD2/CARD15 gene include, without limitation, the JW17 and JW18 variant alleles, which are described below. In one embodiment, a method of the invention is practiced with a fibrostenosis-predisposing allele located in a non-coding region of the NOD2/CARD15 locus which is a JW1, JW15 or JW16 variant allele. In another embodiment, a method of the invention is practiced with a fibrostenosis-predisposing allele located in a promoter region of the NOD2/CARD15 locus which is a JW17 or JW18 variant allele.

As used herein, the term "JW1 variant allele" means a genetic variation at nucleotide 158 of intervening sequence 8 (intron 8) of a NOD2/CARD15 gene. In relation to the AC007728 sequence, the JW1 variant is located at position 128,143. The genetic variation at nucleotide 158 of intron 8 can be, but is not limited to, a single nucleotide substitution, multiple nucleotide substitutions, or a deletion or insertion of one or more nucleotides. The wild type sequence of intron 8 has a cytosine at position 158; as non-limiting examples, a JW1 variant allele can have a cytosine (C) to adenine (A), cytosine to guanine (G), or cytosine to thymine (T) substitution at nucleotide 158 of intron 8. In one embodiment, the JW1 variant allele is a change from a cytosine to a thymine at nucleotide 158 of NOD2/CARD15 intron 8.

As used herein, the term "JW15 variant allele" means a genetic variation in the 3' untranslated region of NOD2/CARD15 at nucleotide position 118,790 of the AC007728 sequence. The genetic variation at nucleotide 118,790 can be, but is not limited to, a single nucleotide substitution, multiple nucleotide substitutions, or a deletion or insertion of one or more nucleotides. The wild type sequence has an adenine (A) at position 118,790; as non-limiting examples, a JW15 variant allele can have an adenine (A) to cytosine (C), adenine to guanine (G), or adenine to thymine (T) substitution at nucleotide 118,790. In one embodiment, the JW15 variant allele is a change from an adenine to a cytosine at nucleotide 118,790.

As used herein, the term "JW16 variant allele" means a genetic variation in the 3' untranslated region of NOD2/CARD15 at nucleotide position 118,031 of the AC007728 sequence. The genetic variation at nucleotide 118,031 can be, but is not limited to, a single nucleotide substitution, multiple nucleotide substitutions, or a deletion or insertion of one or more nucleotides. The wild type sequence has a guanine (G) at position 118,031; as non-limiting examples, a JW16 variant allele can have a guanine (G) to cytosine (C), guanine to adenine (A), or guanine to thymine (T) substitution at nucleotide 118,031. In one embodiment, the JW16 variant allele is a change from a guanine to an adenine at nucleotide 118,031.

As used herein, the term "JW17 variant allele" means a genetic variation in the 5' untranslated region of NOD2/CARD15 at nucleotide position 154,688 of the AC007728 sequence. The genetic variation at nucleotide 154,688 can be, but is not limited to, a single nucleotide substitution,

multiple nucleotide substitutions, or a deletion or insertion of one or more nucleotides. The wild type sequence has a cytosine (C) at position 154,688; as non-limiting examples, a JW17 variant allele can have a  
5 cytosine (C) to guanine (G), cytosine to adenine (A), or cytosine to thymine (T) substitution at nucleotide 154,688. In one embodiment, the JW17 variant allele is a change from a cytosine to a thymine at nucleotide 154,688.

10 As used herein, the term "JW18 variant allele" means a genetic variation in the 5' untranslated region of NOD2/CARD15 at nucleotide position 154,471 of the AC007728 sequence. The genetic variation at nucleotide 154,471 can be, but is not limited to, a single nucleotide substitution,  
15 multiple nucleotide substitutions, or a deletion or insertion of one or more nucleotides. The wild type sequence has a cytosine (C) at position 154,471; as non-limiting examples, a JW18 variant allele can have a cytosine (C) to guanine (G), cytosine to adenine (A), or  
20 cytosine to thymine (T) substitution at nucleotide 154,471. In one embodiment, the JW18 variant allele is a change from a cytosine to a thymine at nucleotide 154,471.

Further provided herein is a method of diagnosing  
25 or predicting susceptibility to a clinical subtype of Crohn's disease characterized by fibrostenosing disease by determining the presence or absence in an individual of at least two fibrostenosis-predisposing alleles linked to a NOD2/CARD15 locus, where the presence of one or more of the  
30 fibrostenosis-predisposing alleles is diagnostic of or predictive of susceptibility to the clinical subtype of Crohn's disease characterized by fibrostenosing disease. In a method of the invention, the at least two

fibrostenosis-predisposing alleles are "2" alleles at SNP 8, SNP 12 or SNP 13.

Further provided herein is a method of diagnosing or predicting susceptibility to a clinical subtype of Crohn's disease characterized by fibrostenosing disease by determining the presence or absence in an individual of at least two fibrostenosis-predisposing alleles linked to a NOD2/CARD15 locus, where the presence of one or more of the fibrostenosis-predisposing alleles is diagnostic of or predictive of susceptibility to the clinical subtype of Crohn's disease characterized by fibrostenosing disease. In one embodiment, the at least two fibrostenosis-predisposing alleles are "2" alleles at SNP 8, SNP 12, or SNP 13. In another embodiment, a method of the invention for diagnosing or predicting susceptibility to a clinical subtype of Crohn's disease characterized by fibrostenosing disease is practiced by determining the presence or absence in the individual of (i) a "2" allele at SNP 8, (ii) a "2" allele at SNP 12, and (iii) a "2" allele at SNP 13, where the presence of one or more of the "2" alleles at SNP 8, SNP 12, and SNP 13 is diagnostic of or predictive of susceptibility to the clinical subtype of Crohn's disease characterized by fibrostenosing disease.

The strength of the association between a fibrostenosing-predisposing allele and Crohn's disease can be characterized by a particular odds ratio such as an odds ratio of at least 2 with a lower 95% confidence interval limit of greater than 1. Such an odds ratio can be, for example, at least 3.0, 4.0, 5.0, 6.0, 7.0, or 8.0 or greater with a lower 95% confidence interval limit of greater than 1, such as an odds ratio of at least 7.4 with a 95%

confidence interval of 1.9-28.9 (see Figure 3). In addition, an odds ratio can be, for example, at least 2.37 with a lower confidence interval limit of 1.26-4.47 (see Figure 3). In one embodiment, the fibrostenosis-predisposing allele is associated with the clinical subtype of Crohn's disease characterized by fibrostenosing disease with an odds ratio of at least 2 and a lower 95% confidence limit greater than 1. Methods for determining an odds ratio are well known in the art (see, for example, Schlesselman et al., Case Control Studies: Design, Conduct and Analysis Oxford University Press, New York (1982)).

In one embodiment, a fibrostenosis-predisposing allele is associated with the fibrostenosing subtype of Crohn's disease with a p value of equal to or less than 0.05. In other embodiments, a fibrostenosis-predisposing allele is associated with the fibrostenosing subtype of Crohn's disease with a p value of equal to or less than 0.1. As used herein, the term "p value" is synonymous with "probability value." As is well known in the art, the expected p value for the association between a random allele and disease is 1.00. A p value of less than about 0.05 indicates that the allele and disease do not appear together by chance but are influenced by positive factors. Generally, the statistical threshold for significance of linkage has been set at a level of allele sharing for which false positives would occur once in twenty genome scans (p=0.05). In particular embodiments, a fibrostenosis-predisposing allele is associated with a clinical subtype of Crohn's disease characterized by fibrostenosis with a p value of equal to or less than 0.1, 0.05, 0.04, 0.03, 0.02, 0.01, 0.009, 0.008, 0.007, 0.006, 0.005, 0.004, 0.003, 0.002 or 0.001, or with a p value of



less than 0.00095, 0.0009, 0.00085, 0.0008 or 0.0005. It is recognized that, in some cases, p values may need to be corrected, for example, to account for factors such as sample size (number of families), genetic heterogeneity, clinical heterogeneity, or analytical approach (parametric or nonparametric method).

A variety of means can be used to determine the presence or absence of a fibrostenosing-predisposing allele in a method of the invention. As an example, enzymatic amplification of nucleic acid from an individual can be conveniently used to obtain nucleic acid for subsequent analysis. The presence or absence of a fibrostenosis-predisposing allele also can be determined directly from the individual's nucleic acid without enzymatic amplification.

Analysis of nucleic acid from an individual, whether amplified or not, can be performed using any of various techniques. Useful techniques include, without limitation, polymerase chain reaction based analysis, sequence analysis and electrophoretic analysis, which can be used alone or in combination. As used herein, the term "nucleic acid" means a polynucleotide such as a single- or double-stranded DNA or RNA molecule including, for example, genomic DNA, cDNA and mRNA. The term nucleic acid encompasses nucleic acid molecules of both natural and synthetic origin as well as molecules of linear, circular or branched configuration representing either the sense or antisense strand, or both, of a native nucleic acid molecule. It is understood that such nucleic acid can be attached to a synthetic material such as a bead or column matrix.

The presence or absence of a fibrostenosing-predisposing allele can involve amplification of an individual's nucleic acid by the polymerase chain reaction. Use of the polymerase chain reaction for the amplification of nucleic acids is well known in the art (see, for example, Mullis et al. (Eds.), The Polymerase Chain Reaction, Birkhäuser, Boston, (1994)). In one embodiment, polymerase chain reaction amplification is performed using one or more fluorescently labeled primers. In another embodiment, polymerase chain reaction amplification is performed using one or more labeled or unlabeled primers that contain a DNA minor groove binder.

Several different primers can be used to amplify an individual's nucleic acid by the polymerase chain reaction. For example, the PCR primers listed in Table 2 (SEQ ID NOS: 37-44) and shown in Figures 5 to 8 can be used to amplify specific regions in the NOD2/CARD15 locus of an individual's nucleic acid. For example, the region surrounding SNP 8 can be amplified using SEQ ID NO: 39 and 40; SNP 12 can be amplified using SEQ ID NOS: 41 and 42, and the region surrounding SNP 13 can be amplified using SEQ ID NOS: 43 and 44. In addition, for example, the region surrounding. As understood by one skilled in the art, additional primers for PCR analysis can be designed based on the sequence flanking the region of interest. As a non-limiting example, a sequence primer can contain about 15 to 30 nucleotides of a sequence upstream or downstream of the region of interest. Such primers are generally designed to have sufficient guanine and cytosine content to attain a high melting temperature which allows for a stable annealing step in the amplification reaction. Several computer

programs, such as Primer Select, are available to aid in the design of PCR primers.

5 A Taqman® allelic discrimination assay available from Applied Biosystems can be useful for determining the presence or absence of a fibrostenosing-predisposing allele. In a Taqman® allelic discrimination assay, a specific, fluorescent, dye-labeled probe for each allele is constructed. The probes contain different fluorescent  
10 reporter dyes such as FAM and VICTM to differentiate the amplification of each allele. In addition, each probe has a quencher dye at one end which quenches fluorescence by fluorescence resonance energy transfer (FRET). During PCR, each probe anneals specifically to complementary sequences  
15 in the nucleic acid from the individual. The 5' nuclease activity of Taq polymerase is used to cleave only probe that hybridize to the allele. Cleavage separates the reporter dye from the quencher dye, resulting in increased fluorescence by the reporter dye. Thus, the fluorescence  
20 signal generated by PCR amplification indicates which alleles are present in the sample. Mismatches between a probe and allele reduce the efficiency of both probe hybridization and cleavage by Taq polymerase, resulting in little to no fluorescent signal. Improved specificity in  
25 allelic discrimination assays can be achieved by conjugating a DNA minor groove binder (MGB) group to a DNA probe as described, for example, in Kutyaev et al., Nuc. Acids Research 28:655-661 (2000). Minor groove binders include, but are not limited to, compounds such as  
30 dihydrocyclopyrroloindole tripeptide (DPI3).

Sequence analysis also can be useful for determining the presence or absence of a fibrostenosing-

predisposing allele in a method of the invention. A fibrostenosing-predisposing allele can be detected by sequence analysis using primers disclosed herein, for example, as the PCR primers listed in Table 2 (SEQ ID NOS: 37-44) and shown in Figures 5 to 8. As understood by one skilled in the art, additional primers for sequence analysis can be designed based on the sequence flanking the region of interest. As a non-limiting example, a sequence primer can contain about 15 to 30 nucleotides of a sequence about 40 to 400 base pairs upstream or downstream of the region of interest. Such primers are generally designed to have sufficient guanine and cytosine content to attain a high melting temperature which allows for a stable annealing step in the sequencing reaction.

The term "sequence analysis," as used herein in reference to one or more nucleic acids, means any manual or automated process by which the order of nucleotides in the nucleic acid is determined. As an example, sequence analysis can be used to determine the nucleotide sequence of a sample of DNA. The term sequence analysis encompasses, without limitation, chemical and enzymatic methods such as dideoxy enzymatic methods including, for example, Maxam-Gilbert and Sanger sequencing as well as variations thereof. The term sequence analysis further encompasses, but is not limited to, capillary array DNA sequencing, which relies on capillary electrophoresis and laser-induced fluorescence detection and can be performed using instruments such as the MegaBACE 1000 or ABI 3700. As additional non-limiting examples, the term sequence analysis encompasses thermal cycle sequencing (Sears et al., Biotechniques 13:626-633 (1992)); solid-phase sequencing (Zimmerman et al., Methods Mol. Cell Biol. 3:39-42 (1992);

and sequencing with mass spectrometry such as matrix-assisted laser desorption/ionization time-of-flight mass spectrometry MALDI-TOF MS (Fu et al., Nature Biotech. 16: 381-384 (1998)). The term sequence analysis also  
5 includes, yet is not limited to, sequencing by hybridization (SBH), which relies on an array of all possible short oligonucleotides to identify a segment of sequences present in an unknown DNA (Chee et al., Science 274:610-614 (1996); Drmanac et al., Science 260:1649-1652 (1993); and Drmanac et  
10 al., Nature Biotech. 16:54-58 (1998)). One skilled in the art understands that these and additional variations are encompassed by the term sequence analysis as defined herein. See, in general, Ausubel et al., *supra*, Chapter 7 and supplement 47.

15

The invention also provides a method of diagnosing or predicting susceptibility to a clinical subtype of Crohn's disease characterized by fibrostenosing disease by determining the presence or absence in an individual of a  
20 fibrostenosis-predisposing allele linked to a NOD2/CARD15 locus, where the presence of the fibrostenosis-predisposing allele is diagnostic of or predictive of susceptibility to the clinical subtype of Crohn's disease characterized by fibrostenosing disease, and where the method includes the  
25 steps of obtaining material containing nucleic acid including the NOD2/CARD15 locus from the individual. As used herein, the term "material" means any biological matter from which nucleic acid can be prepared. As non-limiting examples, the term material encompasses whole blood, plasma,  
30 saliva, cheek swab, or other bodily fluid or tissue that contains nucleic acid. In one embodiment, a method of the invention is practiced with whole blood, which can be obtained readily by non-invasive means and used to prepare

genomic DNA, for example, for enzymatic amplification or automated sequencing. In another embodiment, a method of the invention is practiced with tissue obtained from an individual such as tissue obtained during surgery or biopsy procedures.

Electrophoretic analysis also can be useful in the methods of the invention. Electrophoretic analysis, as used herein in reference to one or more nucleic acids such as amplified fragments, means a process whereby charged molecules are moved through a stationary medium under the influence of an electric field. Electrophoretic migration separates nucleic acids primarily on the basis of their charge, which is in proportion to their size, with smaller molecules migrating more quickly. The term electrophoretic analysis includes, without limitation, analysis using slab gel electrophoresis, such as agarose or polyacrylamide gel electrophoresis, or capillary electrophoresis. Capillary electrophoretic analysis generally occurs inside a small-diameter (50-100-  $\mu$ m) quartz capillary in the presence of high (kilovolt-level) separating voltages with separation times of a few minutes. Using capillary electrophoretic analysis, nucleic acids are conveniently detected by UV absorption or fluorescent labeling, and single-base resolution can be obtained on fragments up to several hundred base pairs. Such methods of electrophoretic analysis, and variations thereof, are well known in the art, as described, for example, in Ausubel et al., Current Protocols in Molecular Biology Chapter 2 (Supplement 45) John Wiley & Sons, Inc. New York (1999).

Restriction fragment length polymorphism (RFLP) analysis also can be useful for determining the presence or

absence of a fibrostenosis-predisposing allele in a method of the invention (Jarcho et al. in Dracopoli et al., Current Protocols in Human Genetics pages 2.7.1-2.7.5, John Wiley & Sons, New York; Innis et al., (Ed.), PCR Protocols, San Diego: Academic Press, Inc. (1990)). As used herein, restriction fragment length polymorphism analysis is any method for distinguishing genetic polymorphisms using a restriction enzyme, which is an endonuclease that catalyzes the degradation of nucleic acid and recognizes a specific base sequence, generally a palindrome or inverted repeat. One skilled in the art understands that the use of RFLP analysis depends upon an enzyme that can differentiate two alleles at a polymorphic site.

Allele-specific oligonucleotide hybridization also can be used to detect the presence or absence of a fibrostenosis-predisposing allele. Allele-specific oligonucleotide hybridization is based on the use of a labeled oligonucleotide probe having a sequence perfectly complementary, for example, to the sequence encompassing a fibrostenosis-predisposing allele. Under appropriate conditions, the allele-specific probe hybridizes to a nucleic acid containing the fibrostenosis-predisposing allele but does not hybridize to the one or more other alleles, which have one or more nucleotide mismatches as compared to the probe. If desired, a second allele-specific oligonucleotide probe that matches an alternate allele also can be used. Similarly, the technique of allele-specific oligonucleotide amplification can be used to selectively amplify, for example, a fibrostenosis-predisposing allele by using an allele-specific oligonucleotide primer that is perfectly complementary to the nucleotide sequence of the fibrostenosis-predisposing allele but which has one or more

mismatches as compared to other alleles (Mullis et al.,  
*supra*, 1994). One skilled in the art understands that the  
one or more nucleotide mismatches that distinguish between  
the fibrostenosis-predisposing allele and one or more other  
5 alleles are often located in the center of an  
allele-specific oligonucleotide primer to be used in  
allele-specific oligonucleotide hybridization. In contrast,  
an allele-specific oligonucleotide primer to be used in PCR  
amplification generally contains the one or more nucleotide  
10 mismatches that distinguish between the disease-associated  
and other alleles at the 3' end of the primer.

A heteroduplex mobility assay (HMA) is another  
well known assay that can be used to detect the presence or  
15 absence of a fibrostenosis-predisposing allele in a method  
of the invention. HMA is useful for detecting the presence  
of a polymorphic sequence since a DNA duplex carrying a  
mismatch has reduced mobility in a polyacrylamide gel  
compared to the mobility of a perfectly base-paired duplex  
20 (Delwart et al., *Science* 262:1257-1261 (1993); White et al.,  
*Genomics* 12:301-306 (1992)).

The technique of single strand conformational  
polymorphism (SSCP) also can be used to detect the presence  
25 or absence of a fibrostenosis-predisposing allele in a  
method of the invention (see Hayashi, *Methods Applic.*  
1:34-38 (1991)). This technique is used to detect mutations  
based on differences in the secondary structure of  
single-strand DNA that produce an altered electrophoretic  
30 mobility upon non-denaturing gel electrophoresis.  
Polymorphic fragments are detected by comparison of the  
electrophoretic pattern of the test fragment to  
corresponding standard fragments containing known alleles.



Denaturing gradient gel electrophoresis (DGGE) also can be used to detect a fibrostenosis-predisposing allele in a method of the invention. In DGGE, double-stranded DNA is electrophoresed in a gel containing an increasing concentration of denaturant; double-stranded fragments made up of mismatched alleles have segments that melt more rapidly, causing such fragments to migrate differently as compared to perfectly complementary sequences (Sheffield et al., "Identifying DNA Polymorphisms by Denaturing Gradient Gel Electrophoresis" in Innis et al., *supra*, 1990).

Other molecular methods useful for determining the presence or absence of a fibrostenosis-predisposing allele are known in the art and useful in the methods of the invention. Other well-known approaches for determining the presence or absence of a fibrostenosis-predisposing allele include, without limitation, automated sequencing and RNAase mismatch techniques (Winter et al., Proc. Natl. Acad. Sci. 82:7575-7579 (1985)). Furthermore, one skilled in the art understands that, where the presence or absence of multiple alleles or a fibrostenosis-predisposing haplotype is to be determined, individual alleles can be detected by any combination of molecular methods. See, in general, Birren et al. (Eds.) Genome Analysis: A Laboratory Manual Volume 1 (Analyzing DNA) New York, Cold Spring Harbor Laboratory Press (1997). In addition, one skilled in the art understands that multiple alleles can be detected in individual reactions or in a single reaction (a "multiplex" assay). In view of the above, one skilled in the art realizes that the methods of the invention for diagnosing or predicting susceptibility to a clinical subtype of Crohn's disease characterized by fibrostenosing disease can be

practiced using one or any combination of the well known assays described above or known in the art.

5 The present invention further provides a method of  
diagnosing or predicting susceptibility to a clinical  
subtype of Crohn's disease characterized by fibrostenosing  
disease by determining the presence or absence in an  
individual of a fibrostenosis-predisposing haplotype, where  
10 the presence of the fibrostenosis-predisposing haplotype is  
diagnostic of or predictive of susceptibility to the  
clinical subtype of Crohn's disease characterized by  
fibrostenosing disease. In one embodiment, the  
fibrostenosis-predisposing haplotype is associated with a  
clinical subtype of Crohn's disease characterized by  
15 fibrostenosing disease with an odds ratio of at least 2 and  
a lower 95% confidence limit greater than 1. In another  
embodiment, the fibrostenosis-predisposing haplotype is  
associated with a clinical subtype of Crohn's disease  
characterized by fibrostenosing disease with an odds ratio  
20 of at least 3, at least 4, at least 5, at least 6, and a  
lower 95% confidence limit greater than 1.

The term "fibrostenosis-predisposing haplotype,"  
as used herein, means a combination of alleles that tend to  
25 be inherited together with the clinical subtype of Crohn's  
disease characterized by fibrostenosing disease more often  
than would be expected according to traditional Mendelian  
genetics. In a method of the invention, the clinical  
subtype of Crohn's disease can be, for example,  
30 characterized by fibrostenosing disease independent of small  
bowel involvement. In one embodiment, the fibrostenosis-  
predisposing haplotype includes at least one allele linked  
to the NOD2/CARD15 locus. In another embodiment, the

fibrostenosis-predisposing haplotype includes a fibrostenosis-predisposing allele. In a further embodiment, a fibrostenosis-predisposing haplotype contains, for example, a variant allele at the NOD2/CARD15 locus. In  
5 another embodiment, the fibrostenosis-predisposing haplotype includes the "2" allele at SNP 8, SNP 12, or SNP 13. In a further embodiment, the fibrostenosis-predisposing haplotype includes the "2" allele at SNP 13. In still further  
10 embodiments, the fibrostenosis-predisposing haplotype includes the "2" allele at SNP 8, SNP 12, and SNP 13. In another embodiment, the fibrostenosis-predisposing haplotype includes a JW1, JW15, JW16, JW17, or JW18 variant allele. One skilled in the art understands that a  
15 fibrostenosis-predisposing haplotype can contain alleles that individually are not significantly associated the fibrostenosing subtype of Crohn's disease, so long as the combination of alleles making up the haplotype tend to be inherited together with the fibrostenosing subtype of  
20 Crohn's disease more often than would be expected according to traditional Mendelian genetics.

The presence or absence of a fibrostenosis-predisposing haplotype can be accomplished using any of the methods described herein above for  
25 determining the presence or absence of a fibrostenosis-predisposing allele. As an example, enzymatic amplification such as polymerase chain reaction amplification, for example, using one or more fluorescently  
30 labeled probes or one or more probes containing a DNA minor groove binder can be useful for determining the presence or absence of a fibrostenosis-predisposing haplotype in a method of the invention.

Antibody based methods also can be useful for determining the presence or absence of a fibrostenosis-predisposing allele or fibrostenosis-predisposing haplotype in a method of the invention. As an example, an antibody that is specifically reactive with a polypeptide or fragment thereof encoded by fibrostenosis-predisposing allele can be used to detect the presence or absence of that allele in an individual. Such an antibody can be, for example, specifically reactive with the truncated version of NOD2/CARD15 generated by a "2" allele at SNP 13 but not reactive with full-length or wild type NOD2/CARD15.

Antibodies useful in the methods of the invention include, without limitation, monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, bifunctional or bispecific antibodies, humanized antibodies, human antibodies, and complementary determining region (CDR)-grafted antibodies, including compounds which include CDR or antigen-binding sequences, which differentially bind to a polypeptide or fragment encoded by a fibrostenosis-predisposing allele but not to other, non-predisposing alleles. Antibody fragments, including Fab, Fab', F(ab')<sub>2</sub>, and Fv, also can be useful in the methods of the invention as can plastic antibodies or molecularly imprinted polymers (MIPs; Haupt and Mosbauch, TIBTech 16:468-475 (1998)). Screening assays to determine differential binding specificity of an antibody are well known in the art (see Harlow et al. (Eds), Antibodies: A Laboratory Manual; Cold Spring Harbor Laboratory; Cold Spring Harbor, N.Y. (1988)).

Antibodies useful in a method of the invention can be produced using any method well known in the art, using a polypeptide, or immunogenic fragment thereof, encoded by a fibrostenosis-predisposing allele. Immunogenic polypeptides or fragments can be isolated from natural sources, from recombinant host cells, or can be chemically synthesized. Methods for synthesizing such peptides are known in the art as described, for example, in Merrifield, J. Amer. Chem. Soc. 85: 2149-2154 (1963); Krstenansky et al., FEBS Lett. 211:10 (1987).

Antibodies that differentially bind to polypeptides encoded by fibrostenosis-predisposing alleles of the invention can be labeled with a detection moiety and used to detect the presence, absence or amount of the encoded polypeptide *in vivo*, *in vitro*, or *in situ*. A moiety, such as a fluorescent molecule, can be linked to an antibody for use in a method of the invention. A moiety such as detection moiety can be linked to an antibody using, for example, carbodiimide conjugation (Bauminger and Wilchek, Meth. Enzymol. 70:151-159 (1980)).

Antibodies that differentially bind to polypeptides encoded by a fibrostenosis-predisposing allele can be used in immunoassays. Immunoassays include, without limitation, radioimmunoassays, enzyme-linked immunosorbent assays (ELISAs) and immunoassays with fluorescently labeled antibodies, which are well known in the art. Antibodies can also be used to detect the presence or absence of a polypeptide of interest in a cell or tissue using immunohistochemistry or other *in situ* assays. Furthermore, cells containing a polypeptide of interest either on the surface of the cell or internally can be detected by an

antibody using assays such as fluorescence activated cell sorting (FACS). One skilled in the art understands that these and other routine assays can be useful for determining the presence or absence of the gene product of a  
5 fibrostenosis-predisposing allele according to a method of the invention.

The methods of the invention optionally include generating a report indicating the presence or absence in a  
10 individual of a fibrostenosis-predisposing allele or fibrostenosis-predisposing haplotype. The methods of the invention also optionally include generating a report indicating the presence or absence in the individual of a clinical subtype of Crohn's disease characterized by  
15 fibrostenosing disease or the risk that an individual has of having or developing the fibrostenosing subtype of Crohn's disease.

A report can be in a variety of forms, for  
20 example, a report can be printed on paper or a report can be an electronic report that is not printed but is transmitted over an electronic medium such as electronic mail or a computer diskette. A report also can be an oral report that indicates the presence or absence in the individual of a  
25 fibrostenosis-predisposing allele or a clinical subtype of Crohn's disease characterized by fibrostenosing disease.

The invention also provides a method of optimizing therapy in an individual by determining the presence or  
30 absence in the individual of a fibrostenosis-predisposing allele linked to a NOD2/CARD15 locus, diagnosing individuals in which the fibrostenosis-predisposing allele is present as having a fibrostenosing subtype of Crohn's disease, and

treating the individual having a fibrostenosing subtype of Crohn's disease based on the diagnosis. Treatment for the fibrostenosing subtype of Crohn's disease currently includes, for example, surgical removal of the affected, strictured part of the bowel. In one embodiment, the presence or absence of a fibrostenosis-predisposing allele is determined in an individual having a known diagnosis of Crohn's disease. In another embodiment, the diagnosis is recorded in the form of a report as described above.

The following examples are intended to illustrate but not limit the present invention.

#### EXAMPLE I

##### SELECTION AND CHARACTERIZATION OF STUDY SUBJECTS

This example describes the clinical characterization of how two cohorts of Crohn's disease patients.

##### A. Selection of Study Subjects

Two cohorts of Crohn's disease patients were consecutively identified from an inflammatory bowel disease referral center (Cedars-Sinai Medical Center Inflammatory Bowel Disease Center). The first cohort of 142 patients, ascertained between 1993-1996 and designated "CD1", was previously described in Vasiliauskas et al., Gut 47:487-496 (2000)). The second cohort of 59 patients (Cohort 2) was collected between 1999-2001 and designated "CD2". A cohort of 175 patients with ulcerative colitis was used as an

inflammatory bowel disease control group. This study, which was reviewed and approved for human subject participation by the Cedars-Sinai Institutional Review Board, involved the collection of clinical, serologic and genetic data from patients consenting to the study.

A diagnosis of Crohn's disease in the cohort patients was defined by the presence of a combination of established features from at least two of the following categories: 1) clinical - perforating or fistulizing disease, obstructive symptoms secondary to small bowel stenosis or stricture; 2) endoscopic - deep linear or serpiginous ulcerations, discrete ulcers in normal-appearing mucosa, cobblestoning, or discontinuous or asymmetric inflammation; 3) radiographic - segmental disease (skip lesions), small bowel or colon strictures, stenosis, or fistula, and; 4) histopathologic - submucosal or transmural inflammation, multiple granulomas, marked focal cryptitis or focal chronic inflammatory infiltration within and between biopsies, or skip lesions including rectal sparing in the absence of local therapy.

To identify clinical features and immunological traits associated with allelic variants of the NOD2/CARD15 gene, the study was designed to analyze two consecutively ascertained cohorts of patients with Crohn's disease. The first cohort was used to explore the relationship of NOD2/CARD15 alleles with an array of clinical and serologic variables, thereby generating hypotheses. The second cohort was then used to confirm the specific hypotheses generated from analysis of the first cohort. To minimize type I error and maximize statistical power, the significance of the associations in the first cohort were permitted to be less



stringent ( $p < 0.1$ ) than in the second cohort ( $p < 0.05$ ). By avoiding a highly stringent correction for the number of variables examined in the first cohort, there was an increased ability to identify specific associations between NOD2/CARD15 allele variants and clinical variables.

#### B. Clinical Characterization of Study Subjects

Patients with Crohn's disease were characterized as having fibrostenosing disease, internal-perforating disease, perianal fistulizing disease or ulcerative colitis (UC)-like disease based on previously described criteria (Gasche et al., Inflammatory Bowel Diseases 6:8-15 (2000); Vasiliasauskas et al., Gut 47:487-496 (2000); Vasiliasauskas et al., Gastroenterology 110:1810-1819 (1996); and Greenstein et al., Gut 29:588-592 (1988)). According to well established criteria, fibrostenosing disease was defined by documented persistent intestinal obstruction or an intestinal resection for an intestinal obstruction. Internal perforating disease was defined as current or previous evidence of entero-enteric or entero-vesicular fistulae, intraabdominal abscesses, or small bowel perforation. Perianal perforating disease was defined by current or previous evidence of either perianal fistulae or abscesses or rectovaginal fistula. UC-like disease was defined by current or previous evidence of left-sided colonic involvement, symptoms of bleeding or urgency, and crypt abscesses on colonic biopsies as previously described. Disease location was classified as small bowel, colon, or both based on one or more endoscopic, radiologic or pathologic studies. A panel of inflammatory bowel disease physicians unaware of the results of serologic and genetic

testing reached a consensus on phenotype based on the clinical data.

Serum ANCA expression and ANCA subtype characterization were performed by fixed neutrophil enzyme-linked immunosorbent assay (ELISA) as previously described (Saxon et al., J. Allergy Clin. Immunol. 86:202-210 (1990)). Briefly, human peripheral blood neutrophils fixed with methanol were reacted with control and coded sera at a 1:100 dilution. Anti-human immunoglobulin G (IgG) (chain-specific) antibody (Jackson ImmunoResearch Labs, Inc.; West Grove, PA) conjugated to alkaline phosphatase was added to label neutrophil-bound antibody, and a colorimetric reaction was performed. Levels were determined relative to a standard consisting of pooled sera obtained from well-characterized pANCA+ UC patients. Results were expressed as ELISA units (EU) per milliliter. ANCA+ sera were further subtyped via indirect immunofluorescent staining to determine the ANCA neutrophil binding pattern, as previously described (Saxon et al., *supra*, 1990). Sera showing the characteristic perinuclear highlighting and losing the characteristic staining pattern when treated with deoxyribonuclease were termed pANCA+ (Vidrich et al., J. Clin. Immunol. 15:293-299 (1995)). For the purposes of this study, patients were considered pANCA+ if they were both positive for ANCA by ELISA and lost perinuclear immunofluorescence staining with deoxyribonuclease treatment.

Sera were analyzed for ASCA expression in a blinded fashion by using a fixed ELISA assay (Vasiliauskas et al., Gut 47:487-496 (2000); and Vermeire et al., Gastroenterology 120:827-833 (2001)). Two patients in the

second cohort did not undergo ASCA testing. High-binding polystyrene microtiter plates were coated with purified phosphopeptidomannans extracted from yeast *S. uvarum*, a subspecies of *S. cerevisiae*. Coded patient sera were  
5 diluted and added to the wells, followed by an enzyme-linked colorimetric reaction. Color development was proportional to concentrations of ASCA antibody present in the sera. Levels were determined and results expressed as EU per milliliter, relative to a standard, which was derived from a  
10 pool of patient sera with well-characterized Crohn's disease found to have reactivity to this antigen. Sera showing ASCA IgG reactivity of >40 EU/mL or IgA reactivity of >20 EU/mL were termed ASCA+. Serological assays were performed at  
15 Cedars-Sinai Medical Center and Prometheus Laboratories (San Diego, CA) using substantially identical methodology. The clinical characteristics of the two Crohn's disease cohorts are shown in Table 1.

TABLE 1			
Clinical Characteristics of Two Crohn's Disease Cohorts			
Clinical characteristics	CD1	CD2	
	n=142	n=59	p
Gender (M/F)	79/63	33/26	0.969
Age at onset	22 (4-67)	22 (2-58)	0.6621
Ethnicity (Jew/Non-Jew)	60/82	23/36	0.668
Disease location (%)			
SB only	19.0	26.4	0.496
Colon only	20.4	20.8	
SB and Colon	60.6	52.8	
Perianal perforating (%)	35.9	28.8	0.332
Internal perforating (%)	47.2	23.7	0.002
Fibrostenosing disease (%)	59.9	30.5	0.001
UC-like (%)	39.4	22.0	0.018
pANCA-positive (%)	19.7	12.5	0.295
ASCA-positive* (%)	57.0	38.6	0.019

## EXAMPLE II

PATIENTS WITH CROHN'S DISEASE HAVE AN INCREASED FREQUENCY OF  
RARE ALLELIC VARIANTS OF NOD2/CARD15

This example describes the association of a "2" allele at SNP 8, SNP 12, or SNP 13 within the NOD2/CARD15 locus (rare allelic variants of NOD2/CARD15) with Crohn's disease in a North American population.

In order to determine whether the North American Crohn's disease patient populations in Cohorts 1 and 2 expressed allelic variants of NOD2/CARD15, Cohort 1 (hypothesis-generating) and Cohort 2 (hypothesis-confirming)

were genotyped for the rare allelic variant of SNP 8 (R675W), SNP 12 (G881R) and SNP 13 (3020insC). A cohort of ulcerative colitis patients was used for comparison.

5                    Genotyping was performed using a genotyping assay  
employing 5'-exonuclease technology, the TaqMan MGB™ assay  
(PE Biosystems; Foster City, CA). Primers were designed  
using the software PrimerExpress 1.5™ (PE Biosystems) and  
sequence information found in dbSNP for NOD2/CARD15 SNP 5,  
10    8, 12, and 13. The MGB™ design adds a "minor groove binder"  
to the 3' end of the TaqMan™ probes, thereby increasing the  
binding temperature of the probe and enabling the use of  
shorter probes than in conventional TaqMan™ assays (Kutyavin  
et al., Nucleic Acids Res. 25:3718-3723 (1997)). This has  
15    the effect of increasing the discrimination between the  
alleles in the assay (Kutyavin et al., Nucleic Acids Res.  
28:655-661 (2000)). Assays were performed following the  
manufacturer's recommendations (PE Biosystems bulletin  
4317594) in an ABI 7900 instrument. Genotyping was  
20    performed blinded to clinical status of the subjects.  
Primers and probes used in the genotyping assay are shown in  
Tables 2 and 3. ✓

Table 2 Primers Used in Taqman MGB™ Assay for SNPs 5, 8, 12 and 13				
SNP Primer	Forward Primer	Reverse Primer	SEQ ID NO	
5	5'GGTGGCTGGGC TCTTCT 3'	5'CTCGCTTCCTCAGTACCTAT GATG 3'	for 37 rev 38	
8	5'CTGGCTGAGTG CCAGACATCT 3'	5'GGCGGGATGGAGTGGAA 3'	for 39 rev 40	
12	5'CCACCTCAAGC TCTGGTGATC 3'	5'GTTGACTCTTTTGGCCTTTT CAG 3'	for 41 rev 42	
13	5'CCTTACCAGAC TTCCAGGATGGT 3'	5'TGTCCAATAACTGCATCACC TACCT 3'	for 43 rev 44	

Table 3 TAQMAN PROBES		
Allele detected	Probe sequence	Seq ID NO
SNP5 wild type allele ("1")	6FAM-CATGGCTGGACCC-MGBNFQ	45
SNP5 variant allele ("2")	TET-CATGGCTGGATCC-MGBNFQ	46
SNP8 wild type allele ("1")	6FAM-TGCTCCGGCGCCA-MGBNFQ	47
SNP8 variant allele ("2")	TET-CTGCTCTGGCGCCA-MGBNFQ	48
SNP12 wild type allele ("1")	6FAM-CTCTGTTGCCCCAGAA- MGBNFQ	49
SNP12 wild type allele ("1")	TET-CTCTGTTGCGCCAGA-MGBNFQ	50
SNP13 wild type allele ("1")	TET-CTTTCAAGGGCCTGC-MGBNFQ	51
SNP13 variant allele ("2")	6FAM-CCTTTCAAGGGGCCT-MGBNFQ	52
JW1 wild type allele	6FAM-AAGACTCGAGTGTCTC- MGBNFQ	53
JW1 variant	VIC-AGACTCAAGTGTCTC-MGBNFQ	54

As shown in Table 4, each of three rare allelic variants of NOD2/CARD15 (a "2" allele at SNP 8, SNP 12, or SNP 13) was significantly more frequent in patients with Crohn's disease compared with ulcerative colitis. In addition, as can be seen in Table 4, the frequency of each of the NOD2/CARD15 rare allelic variants was similar in each cohort of Crohn's disease patients, supporting the combined use of the two cohorts. The overall frequency of any of the three NOD2/CARD15 rare allelic variants was 35% in Crohn's disease patients compared with 11% in ulcerative colitis patients ( $p=0.001$ ).

Within the combined Crohn's disease cohort, the frequency of homozygotes with a "2" allele at SNP 13 (3020insC) and compound rare allelic heterozygotes was 1% and 4%, respectively, while none of the ulcerative colitis patients had such a genotype. These results demonstrate that rare allelic variants of NOD2/CARD15 are associated with Crohn's disease across diverse geographic and ethnically-defined patient populations.



Table 4  
Frequency of NOD2/CARD15 Rare Allelic Variants in CD and UC Patient Populations

Allelic variants	UC (n=175)	CD1 (n=142)	CD2 (n=59)	Combined CD1 and CD2 (n=201)	p (UC vs. Combined CD)
R675W (SNP 8 "2" allele)	5.7%	16.9%	15.3%	16.4%	0.001
G881R (SNP 12 "2" allele)	1.7%	12.0%	10.2%	11.4%	0.0001
3020insC (SNP 13 "2" allele)	3.4%	11.3%	11.9%	11.4%	0.004
Carriage of any allelic variant	10.9%	36.6%	32.2%	35.3%	0.001

## EXAMPLE III

RARE VARIANT ALLELES IN THE NOD2/CARD15 LOCUS ARE ASSOCIATED  
WITH THE FIBROSTENOSING SUBTYPE OF CROHN'S DISEASE IN  
COHORT 1

5

This example demonstrates that a "2" allele at SNP 8, SNP 12, or SNP 13 is significantly associated with fibrostenosing disease in Cohort 1.

10

15

20

25

30

Patients with Crohn's disease express diverse clinical phenotypes that can be due to differences in underlying genetic factors. In order to determine whether rare variant alleles at the NOD2/CARD15 locus were associated with specific Crohn's disease-related clinical phenotypes or disease-related serum immune markers, univariate analysis was performed. The univariate analysis evaluated the association between NOD2/CARD15 allelic variants at SNP 8, SNP 12, or SNP 13 and predefined clinical characteristics, including age of onset, disease location, and disease phenotype (fibrostenosing disease, internal-perforating disease, perianal fistulizing disease or ulcerative colitis-like disease). The association between NOD2/CARD15 allelic variants and expression of the serum immune markers ASCA and pANCA was also tested.

As shown in Table 5, univariate analysis indicated that a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus was significantly associated with fibrostenosing disease in Cohort 1 ( $p=0.049$ ) for the three allelic variants combined. A positive association at a less stringent significance level ( $p<0.1$ ) was also observed with small bowel involvement and younger age of onset, and a

negative association was observed with UC-like disease in this cohort. With respect to serologic markers, patients with the "2" allele at SNP 13 were more likely to express ASCA ( $p=0.053$ ). These results demonstrate that a "2" allele at SNP 8, SNP 12, or SNP 13 is significantly associated with fibrostenosing disease in Cohort 1. .

Statistical analysis was performed using SAS computer software (Version 6.10; SAS Institute, Inc.; Cary, NC). Quantitative variables were described as medians with a range. Nonparametric statistical tests were used to test differences in quantitative variables between two groups. A Chi-square test or Fisher's exact test (when the expected number was less than 5) was used to evaluate associations between carriers and non-carriers of the rare alleles or between genotypes and categorical variables, such as type of IBD, disease location, disease behavior, and antibody positivity. In addition, a Mantel Haenszel stratified association test was performed for all genotype and phenotype associations by controlling for potential confounding effect due to ethnic variation. This stratified association test was also used to evaluate whether the small bowel involvement and fibrostenosing disease were independently associated with NOD2/CARD15 variants (see Example IX below).

Table 5  
Relationship of NOD2/CARD15 Rare Variant Alleles and Clinical  
Phenotypes of Crohn's Disease in Cohort 1

Clinical phenotypes		Qualitative Trait %NOD2 variant carriers			
		R675W (SNP 8)	G881R (SNP 12)	3020insC (SNP 13)	Carriage of any allelic variant
Small bowel involvement	n				
	yes 113	19.47%	11.50%	14.16%	40.71%
	no 29	6.90%	13.79%	0.00%	20.69%
p value		0.081	0.494	0.04	0.063
Perianal perforating	yes 51	11.76%	11.76%	13.73%	35.29%
	no 91	19.78%	12.09%	9.89%	37.36%
	p value	0.248	0.839	0.547	0.747
Internal perforating	yes 67	13.43%	16.42%	11.94%	37.31%
	no 75	20.00%	8.00%	10.67%	36.00%
	p value	0.346	0.178	0.91	0.96
Fibrostening	yes 85	18.82%	14.12%	15.29%	43.53%
	no 57	14.04%	8.77%	5.26%	26.32%
	p value	0.389	0.458	0.084	0.049
UC-like	yes 56	17.86%	10.71%	5.36%	30.36%
	no 86	16.28%	12.79%	15.12%	40.70%
	p value	0.822	0.736	0.076	0.22
pANCA positive	yes 28	17.86%	14.29%	7.14%	32.14%
	no 114	16.67%	11.40%	12.28%	37.72%
	p value	0.82	0.793	0.394	0.529
ASCA positive	yes 81	18.52%	9.88%	16.05%	38.27%
	no 61	14.75%	14.75%	4.92%	34.43%
	p value	0.467	0.234	0.053	0.744
		Quantitative Trait median (range)			
		R675W (SNP 8)	G881R (SNP 12)	3020insC (SNP 13)	Carriage of any allelic variant
Age of onset carrier of NOD2 variant	yes	22 (6-67)	22 (4-62)	19 (10-50)	20 (4-67)
	no	22 (4-63)	22 (4-67)	22 (4-67)	22 (4-63)
	p	0.715	0.937	0.074	0.238

## EXAMPLE IV

RARE VARIANT ALLELES IN THE NOD2/CARD15 LOCUS ARE ASSOCIATED  
WITH THE FIBROSTENOSING SUBTYPE OF CROHN'S DISEASE IN  
COHORT 2

5

This example demonstrates that a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus is significantly associated with fibrostenosing disease in Cohort 2.

10

The results obtained with Cohort 1 in Example III indicated that NOD2/CARD15 rare variant alleles were positively associated with fibrostenosing Crohn's disease, small bowel involvement, ASCA positivity, and younger age of onset, and negatively associated with UC-like disease (see Table 5). These hypotheses were further tested using Cohort 2; the results are shown in Table 6. As with Cohort 1, Cohort 2 demonstrated a significant association between a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus and fibrostenosing disease ( $p=0.002$ , with Bonferroni correction  $p=0.01$ ; see Table 6). These results indicate that a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus is significantly associated with fibrostenosing disease in Cohort 2.

20

Table 6  
Relationship of NOD2/CARD15 Rare Variant Alleles and Clinical  
Phenotypes of Crohn's Disease in Cohort 2

Clinical phenotypes		n	Qualitative Trait %NOD2 variant carriers			
			R675W (SNP 8)	G881R (SNP 12)	3020insC (SNP 13)	Carriage of any allelic variant
Fibrostenosing	yes	18	22.22%	22.22%	27.78%	61.11%
	no	41	12.20%	4.88%	4.88%	19.51%
	p		0.315	0.048	0.018	0.002
Small bowel involvement	yes	42	19.05%	9.52%	14.29%	35.71%
	no	17	5.88%	11.76%	5.88%	23.53%
	p		0.22	0.828	0.288	0.354
ASCA positive	yes	22	9.09%	13.64%	13.64%	31.82%
	no	35	17.14%	8.57%	11.43%	31.43%
	p		0.4	0.542	0.735	0.956

Quantitative Trait median (range)						
Age of onset carrier of NOD2 variant		n	R675W (SNP 8)	G881R (SNP 12)	3020insC (SNP 13)	Carriage of any allelic variant
			27 (10-58)	26 (7-33)	17 (13-35)	22 (7-58)
			19 (2-55)	20 (2-58)	24 (2-58)	22 (2-55)
p			0.332	0.9	0.566	0.981

## EXAMPLE V

RARE VARIANT ALLELES IN THE NOD2/CARD15 LOCUS ARE ASSOCIATED  
WITH THE FIBROSTENOSING SUBTYPE OF CROHN'S DISEASE IN A  
COMBINED COHORT REPRESENTING COHORTS 1 AND 2

5

This example demonstrates that a "2" allele at SNP 8, SNP 12, or SNP 13 is significantly associated with fibrostenosing disease in a combined cohort representing cohorts 1 and 2.

10

As seen in Figure 2, the association between a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus (NOD2 variant carrier) and fibrostenosing disease was even more significant ( $p=0.001$ ) when the two cohorts were analyzed together than when analyzed separately.

15

The association between a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus (NOD2 variant carrier) and fibrostenosing disease was observed in both Jewish and non-Jewish individuals. Approximately 46% of Crohn's disease patients with fibrostenosing disease (Jews 52% vs. non-Jews 42%) had at least one of these rare alleles compared with only 23% (Jews 21.6% vs. non-Jews 25%) of Crohn's disease patients without fibrostenosing disease (Odds ratio, 2.8; 95% Confidence interval, 1.56-5.18). Of the three rare variant alleles, the frameshift mutation 3020insC (a "2" allele at SNP 13) demonstrated the greatest association with fibrostenosing disease (47% vs. 17%,  $p=0.006$  for cohorts combined). These results indicate that rare variant alleles in the NOD2/CARD15 locus are associated with the fibrostenosing subtype of Crohn's disease in a combined cohort representing Cohorts 1 and 2.

30

## EXAMPLE VI

CROHN'S DISEASE PATIENTS WITH HOMOZYGOUS MUTATIONS OR  
COMPOUND HETEROZYGOUS MUTATIONS IN NOD2/CARD15

5           This example describes the increased risk of  
fibrostenosing disease in Crohn's disease patients carrying  
homozygous mutations or compound heterozygous mutations in  
NOD2/CARD15 locus.

10           As shown in Figure 3, compared with patients who  
were not carriers of NOD2/CARD15 mutations at SNP 8, SNP 12  
or SNP 13, patients who were carriers of two mutations in  
NOD2/CARD15 were significantly more likely to have  
15           fibrostenosing disease (85% vs. 43%; odds ratio 7.4; 95%  
confidence interval 1.9-28.9,  $p=0.004$ ). Patients who were  
carriers of a single NOD2/CARD15 mutation were also  
significantly more likely to have fibrostenosing disease  
when compared with patients who were not carriers of these  
NOD2/CARD15 mutations (64% vs. 43%; Odds ratio 2.37; 95%  
20           Confidence interval 1.26-4.47;  $p=0.008$ ). These results  
indicate that Crohn's disease patients with homozygous  
mutations or compound heterozygous mutations in NOD2/CARD15  
have an increased risk of fibrostenosing disease.

25

## EXAMPLE VII

FIBROSTENOSING DISEASE ONLY COMPARED TO FIBROSTENOSING AND  
PERFORATING DISEASE

30           This example describes the association of the "2"  
allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus  
in patients with fibrostenosing disease compared to patients  
with fibrostenosing and perforating disease.



Fibrostenosing and perforating disease often occur in the same patient. Patients with fibrostenosing disease can be characterized as i) having only fibrostenosing disease or ii) having both fibrostenosing and perforating disease. In order to address these phenotypes individually, patients in each of the two cohorts were separated by the presence of fibrostenosing disease with perforating complications (Fib + perf) or without perforating complications (Fib only) and compared with patients with perforating complications without evidence of fibrostenosis (Perf only). The percentage of patients having only fibrostenosing disease that carried a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus was 48.3%, which was similar to that seen in patients with both fibrostenosing and perforating complications (46.0%;  $p=0.8$ ). As seen in Figure 3, when patients with fibrostenosing disease were compared with those patients described as having perforating disease only (perianal or internal), the frequency of the "2" allele at SNP 8, SNP 12 or SNP 13 of the NOD2/CARD15 locus in patients with fibrostenosing disease (with or without perforating complications) was significantly greater than that seen in patients with only perforating complications (46.6% versus 18.6%;  $p=0.002$ ).

#### EXAMPLE VIII

##### MULTIVARIANT ANALYSIS OF THE COMBINED PATIENT COHORT

This example demonstrates that fibrostenosing disease is independently associated with a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus.

For multivariant analysis, all variables with at least borderline significance ( $p < 0.1$ ) in either cohort were tested simultaneously for their association with a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus by using logistic regression. As shown in Table 7, the clinical phenotype of fibrostenosing disease was significantly associated ( $p < 0.05$ ) with these rare alleles at NOD2/CARD15 locus (OR 2.8; 95% CI, 1.3-6.0). These results confirm that fibrostenosing disease is independently associated with a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus.

Table 7  
Multivariate Analysis in the Combined  
Cohort for 5 Phenotypic Variables

Clinical phenotypes	OR	95% CI	P
Fibrostenosing disease	2.8	1.3-6.0	0.011
Small bowel involvement	1.3	0.5-3.4	0.561
UC-like	0.9	0.4-1.7	0.658
ASCA positive	0.7	0.3-1.3	0.250
Age of onset	1.0	0.9-1.0	0.874

#### EXAMPLE IX

##### FIBROSTENOSING DISEASE AND SMALL-BOWEL INVOLVEMENT

This example demonstrates that the association between fibrostenosing disease and a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus is independent of small-bowel involvement.

Because fibrostenosing disease is more likely to occur in patients with small-bowel involvement, patients were stratified on the basis of small-bowel involvement to analyze whether the association between fibrostenosing disease and NOD2/CARD15 variant alleles was a primary association. Among patients with small-bowel involvement, 26.4% of patients who did not have fibrostenosing disease (n=53) had a "2" allele at SNP 8, SNP 12, or SNP 13, whereas 46.1% of patients who had fibrostenosing disease (n=102) had a "2" allele at SNP 8, SNP 12, or SNP 13 (p=0.017). A similar trend was observed among patients without small-bowel involvement (p=0.05), and the combined analysis conditioning on small-bowel involvement yielded a significance level of 0.009.

After controlling for fibrostenosing disease, small-bowel involvement was not associated with a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus (p=0.63). This result agrees with the results from logistic regression analysis (see Example VIII) and indicates that the association between fibrostenosing disease and a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus is independent of small-bowel involvement. These results further indicate that the observed small-bowel association with a "2" allele at SNP 8, SNP 12, or SNP 13 of the NOD2/CARD15 locus is secondary to the presence of fibrostenosing disease.

All journal article, reference, and patent citations provided above, in parentheses or otherwise, whether previously stated or not, are incorporated herein by reference.

Although the invention has been described with reference to the examples above, it should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the  
5 invention is limited only by the following claims.

We claim:

1. A method of diagnosing or predicting susceptibility to a clinical subtype of Crohn's disease characterized by fibrostenosing disease, comprising  
5 determining the presence or absence in an individual of a fibrostenosis-predisposing allele linked to a NOD2/CARD15 locus,  
wherein the presence of said  
10 fibrostenosis-predisposing allele is diagnostic of or predictive of susceptibility to the clinical subtype of Crohn's disease characterized by fibrostenosing disease.
2. The method of claim 1, wherein said clinical  
15 subtype of Crohn's disease is characterized by fibrostenosing disease independent of small bowel involvement.
3. The method of claim 1, wherein said  
20 fibrostenosis-predisposing allele is located within said NOD2/CARD15 locus.
4. The method of claim 3, wherein NF-kappa B activation by a NOD2/CARD15 polypeptide encoded by said  
25 fibrostenosis-predisposing allele is reduced as compared to NF-kappa B activation by a wild-type NOD2/CARD15 polypeptide.
5. The method of claim 3, wherein said  
30 fibrostenosis-predisposing allele is located in a coding region of said NOD2/CARD15 locus.

6. The method of claim 5, wherein said fibrostenosis-predisposing allele is located in a region encoding residues 744 to 1020 of NOD2/CARD15.

5           7. The method of claim 5, wherein said fibrostenosis-predisposing allele is a "2" allele at a SNP selected from SNP 8, SNP 12, and SNP 13.

10           8. The method of claim 7, wherein said fibrostenosis-predisposing allele is a "2" allele at SNP 13.

15           9. The method of claim 3, wherein said fibrostenosis-predisposing allele is located in a non-coding region of said NOD2/CARD15 locus.

          10. The method of claim 9, wherein said fibrostenosis-predisposing allele is selected from a JW1, JW15, and JW16 variant allele.

20           11. The method of claim 9, wherein said fibrostenosis-predisposing allele is located in a promoter region of said NOD2/CARD15 locus.

25           12. The method of claim 11, wherein said fibrostenosis-predisposing allele is an allele selected from a JW17 and JW18 variant allele.

13. The method of claim 1, comprising determining the presence or absence in said individual of at least two fibrostenosis-predisposing alleles linked to a NOD2/CARD15 locus,

5 wherein the presence of one or more of said fibrostenosis-predisposing alleles is diagnostic of or predictive of susceptibility to the clinical subtype of Crohn's disease characterized by fibrostenosing disease.

10 14. The method of claim 13, wherein said at least two fibrostenosis-predisposing alleles are "2" alleles at a SNP selected from SNP 8, SNP 12, and SNP 13.

15 15. The method of claim 14, comprising determining the presence or absence in said individual of

(i) a "2" allele at SNP 8,

(ii) a "2" allele at SNP 12, and

20 (iii) a "2" allele at SNP 13,

25 wherein the presence of one or more of said "2" alleles at SNP 8, SNP 12, and SNP 13 is diagnostic of or predictive of susceptibility to the clinical subtype of Crohn's disease characterized by fibrostenosing disease.

30 16. The method of claim 1, wherein said fibrostenosis-predisposing allele is associated with said clinical subtype of Crohn's disease characterized by fibrostenosing disease with an odds ratio of at least 2 and a lower 95% confidence limit greater than 1.

17. The method of claim 1, further comprising generating a report indicating the presence or absence in said individual of said fibrostenosis-predisposing allele.

5 18. The method of claim 1, further comprising generating a report indicating the presence or absence in said individual of said clinical subtype of Crohn's disease characterized by fibrostenosing disease.

10 19. The method of claim 1, wherein determining the presence or absence of said fibrostenosis-predisposing allele comprises enzymatic amplification of nucleic acid from said individual.

15 20. The method of claim 19, wherein said amplification is polymerase chain reaction amplification.

20 21. The method of claim 20, wherein said polymerase chain reaction amplification is performed using one or more fluorescently labeled probes.

25 22. The method of claim 20, wherein said polymerase chain reaction amplification is performed using one or more probes comprising a DNA minor groove binder.



23. A method of optimizing therapy in an individual, comprising

5 (a) determining the presence or absence in said individual of a fibrostenosis-predisposing allele linked to a NOD2/CARD15 locus,

10 (b) diagnosing individuals in which said fibrostenosis-predisposing allele is present as having a fibrostenosing subtype of Crohn's disease, and

(c) treating said individual having a fibrostenosing subtype of Crohn's disease based on said diagnosis.

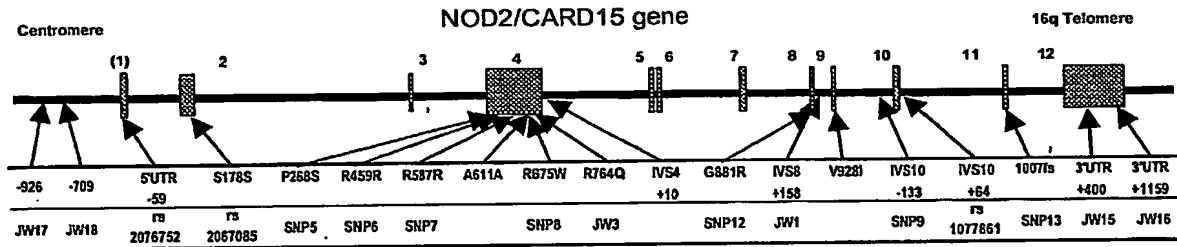
ABSTRACT OF THE DISCLOSURE  
METHODS OF USING A NOD2/CARD15 HAPLOTYPE TO  
DIAGNOSE CROHN'S DISEASE

5           The present invention provides a method of  
diagnosing or predicting susceptibility to a clinical  
subtype of Crohn's disease characterized by fibrostenosing  
disease by determining the presence or absence in an  
individual of a fibrostenosis-predisposing allele linked to  
10 a NOD2/CARD15 locus, where the presence of the  
fibrostenosis-predisposing allele is diagnostic of or  
predictive of susceptibility to the clinical subtype of  
Crohn's disease characterized by fibrostenosing disease. In  
a method of the invention, the clinical subtype of Crohn's  
15 disease can be, for example, characterized by fibrostenosing  
disease independent of small bowel involvement. The  
invention also provides a method of optimizing therapy in an  
individual by determining the presence or absence in the  
individual of a fibrostenosis-predisposing allele linked to  
20 a NOD2/CARD15 locus, diagnosing individuals in which the  
fibrostenosis-predisposing allele is present as having a  
fibrostenosing subtype of Crohn's disease, and treating the  
individual having a fibrostenosing subtype of Crohn's  
disease based on the diagnosis.

25

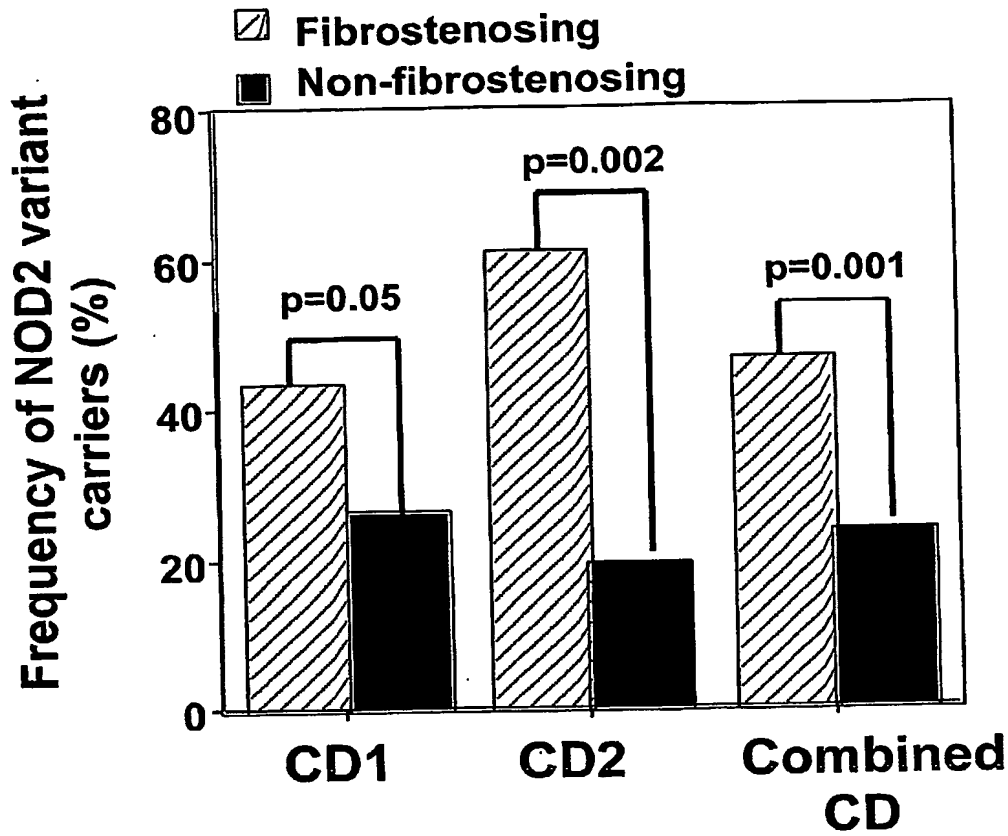
1/15

Figure 1



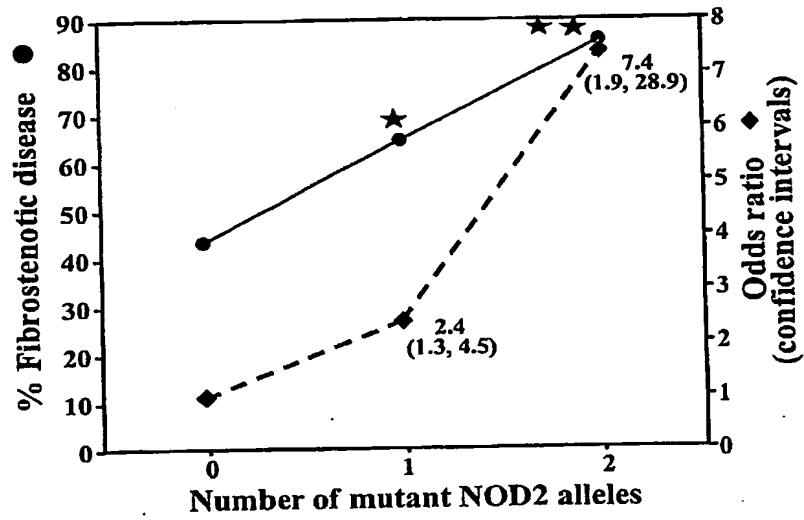
2/15

Figure 2



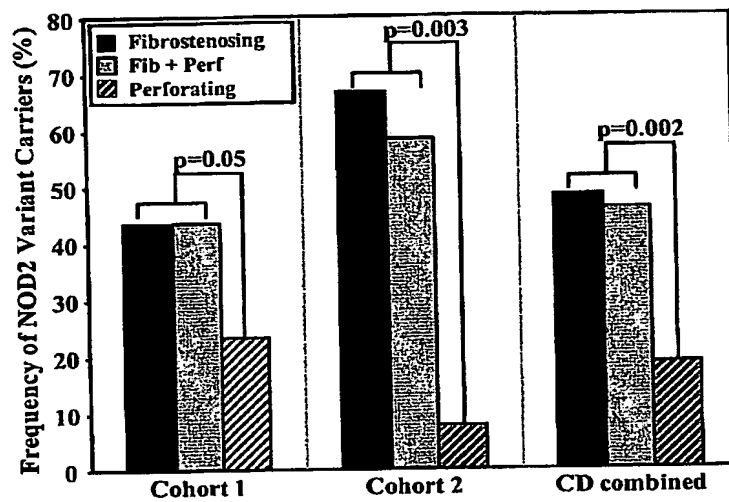
3/15

Figure 3



4/15

Figure 4



5/15

FIGURE 5  
SNP 8

```
5' ACCATCAGAT CACAGCAGCC TTGCTGGCAG GGCCTGTGTC CCGGCACCAC 50
3' TCGAAGTCTA GTGTGCTCGG AAGGACGGTC CCGACACAG GGCCTGTGTC

TGGGCGCTC TGGCTCAGTC CACACATCT CACAGGCCC TGTCCGCTCG 100
ACCCCGGACG ACGACTACAC GGTCTGTAGA CTCTTCGGG .acgaggccgc

CGTGGCTGTG CCGGCTGGT GTCTGGCCCG CAGCCTCCCG AAGCACTTCC 150
GTCACGACAC GGGCCGACCA CAGACCGGCG GTCCGAGCG TTGCTGAGG

ACTCCATCCG GGCAGCTGCA CCGGGTGAGG CCAACAGCGT GCATGCCATG 200
TACGGTAGCG CCGTCCAGCT GGCCTACTCC GGTCTGTCCA CGTACGGTAC

CCCGGGTTCA TTGCGTCTAT CCGGAGCCTG TACGAGATGC AGGAGGAGCG 250
GGGCGCAAGT AGACCGAGTA GGGCTGGAC ATGCTCTACG TCTCTCTGCG

GCTGGCTCGG AAGCTTCAC GTGGCTGTA TGTGGGGAC CTCAGTTGA 300
CGACCGAGCC TTCCGACGTG CACCGCACTT ACAACCGTG GAGTTCACT

CATTTTGCAG TTGGGGCCCC ACTGAGTGTG CTGCCCCGGC CTTTGTGCTG 350
GTAAAACGTC ACACCGGGG TGAATCACAC GACGGGACCG GAAACAGGAC

CAGCACTTC GCGGGCCCGT GCGGCTGCAG CTGGACTACA ACTCTGTGG 400
GTGTTGGAGG CCGCGGGCA CCGGACGGTC GACCTGATGT TGACACACCC

TGACATTGGC GTGGAGGAGC TCTGCGCTG CTTTGGTGTG TCGAAGGCTC 450
ACTGTACCG CACCTGTGTC ACCACCGAAC GGAACACAG ACGTTCCCG

TGATGTGAGT GTTACTGGGC ATTGCTGTTC AGGTATGGGG CAGC 3' SEQ ID NO 1 494
ACATCACTCA CAATGACCGG TACGACAGG TCCATACCCC CTCG 5' SEQ ID NO 2
```

6/15

Figure 6

SNP 12

```
5' ATCAAACCC TGAGAGGACA AGGGACATTT CCAAGTCACC CAGAAAGACT 50
3' TAGTTTGGG ACTCTCTGT TCCCTGTAA GGTTCAGTGG GTCTTTCGA

CGAGTGTCCT CTCTTGAAAT CCAATGGTCT TTTTTCCTTA CTCCATTGCC 100
GCTCACAGGA GAGAACTTTA GGTTACCAGA AAAAAGGAAT GAGGTAACGG

TAACATTGTG GGGTAGAAAT AAAGTTCAAA GACCTTCAGA ACTGGCCCCA 150
ATTGTAAACAC CCCATCTTTA TTTCAAGTTT CTGGAAGTCT TGACCGGGGT

GCTCCTCCCT CTTCACCTGA TCTCCCCAAG AAAACTGCAG GATAGACTCT 200
CGAGGAGGGA GAAGTGGACT AGAGGGGTTC TTTTGACGTC CTATCTGAGA

GAAGCTTACC TGAGCCACCT CAAGCTCTGG TGATCACCA AGGCTTCAGC 250
CTTCGAATGG ACTCGGTGGA GTTCGAGAAC ACTAGTGGGT TCGAAGTGG

CAGGGCCTGG GCCCCCTCGT CACCCActct gttgccccag aaTCTGAAAA 300
GTCCCGGACC CGGGGGAGCA GTGGGTgaga caacgggggtc ttACACTTTT

GGCCAAAAGA GTCAACAGAC AGTGTCAGTG AGTACCTGAT ATGTGTTCTA 350
COGGTTTCT CAGTTCTCTG TCACAGTCAC TCATGGACTA TACACAAGAT

GACATGAACT AACAGTCTTC CTCCCTCTGC AGTCCAGCC AGAGGGGCAG 400
CTGFACTTGA TTGTACGAG GAGGGAGAOG TCAGGGTGG TCTCCCCGTC

GACCACTCAA TCCCAGAGTG GCCTCACTGG GGCTCTGGT CCCAGCAAAG 450
CTGGTGAGTT AGGGTCTCAC OGGAGTGACC CCGAGGACCA GGGTCGTTTC

TGGACCTGCC TCCATCTTTT GGGTGGGATG GCCAACTTA ACCCAAGAGT 500
AAGTGAACG AGGTAGAAAA CCCACCCTAC CGGTTTGAAT TGGGTCTCA

TTTCAGTGGC TTTACATTAC AGACTTAGAG AATAGTAGAG 3'-SEQ ID NO 3 540
AAAGTCACCG AAATGTAATG TCTGAATCTC TTATCATCTC 5'-SEQ ID NO 4
```



7/15

FIGURE 7

SNP 13

```
5' TTTAAAAATG AAATCATTGC TCCCTACTTA AAGAGGTAAA CACTTCTTTC 50
3' AAATTTTTCAC TTTAGTAAAG AGGATGAAT TTCTCCATTT CTGAAGAAAG

TTACACAGAG AATCAGATCC TTCACATCCA GAATCATTCT CACTGAATGT 100
AATCTGTCTC TTAGTCTAGG AAGTGTACGT CTTAGTAAAG GTGACTTACA

CAGAATCAGA AGGATCCTC AAAATTCTGC CATTCTCTC TCCGTCACC 150
GTCTTAGTCT TCCCTAGGAG TTTTAAGACG GTAAGGAGAG AGGGCAGTGC

CCATTTTACA CATAGAAAAA CTGAGGTTCG GAGAGCTAAA ACAGGCTTGC 200
GGTAAAATGT CTATCTTTT CACTCCAAGC CTCCTGATTT TGTCGGGACG

CCAGGGCCT TACCAGACTT CCAGGATGGT GTCATTCctt tcaagggggcc 250
GGTCCCGGA ATGGTCTGAA GGTCCTACCA CAGTAAggaa agttccccgg

tgcAGGAGGG CTCTGCCCC TAGGTAGGTG ATGCAGTTAT TGGACAACCT 300
acgTCCTOCC GAAGACGGGG ATCCATCCAC TACGTCAATA ACCTGTTCGA

GGAAAAGAAG ATACAATGGT GAGCTTCAAG CATTCTTGGT TTCTCTCTTG 350
CCTTTCTTC TATGTTACCA CTCGAAGTTC CTAAGAACCA AAAGGACAAC

AAACTGTCCA GTTAAAGAGA CTGCAGGAGT TAGCCAGTCT ACTGAAGCCC 400
TTTGACAGGT CAATTCTCT CAGCTCTCA ATCGGTCAGA TGACTIONGG

ACCTGTCCCT TAGACACATC CTGCTCATGT CTGAGATTCC CAATGAGCTC 450
TGGACAGGCA ATCTGTGTAG GACGAGTACA GACTCTAAGG GTTACTOCAG

ATCAACAAAG GCTCAGTACC ATCAGTGAAA TGTAAOOGTC TCTCTTCCAT 500
TAGTTGTTTC CGAGTCATGG TAGTCACTTT ACATTGCCAG AGAGAAGGTA

TCACTAGATG AGTTTATCAA ATTAAGTAGC CACTCCCTTA G3'-SEQ ID NO 5 541
AGTGATCTAC TCAATAGTT TAATTCATCG GTGAGCGAAT C5'-SEQ ID NO 6
```

8/15

FIGURE 8

SNP 5

5' AACAGCAGTG CTCAAAGAGT AGAGTCCGCA CAGAGAGTGG TTTCGCCATG 50  
3' TTGTGCTCAC GAGTTTCTCA TCTCAGGCGT GTCTCTCACC AAACCGGTAC  
  
CACTGCAGCT GCGGGCAGCT GAATGGGAAG ACAAAGAGAA ATTCTCGAA 100  
GTGACGTGCA CGGCGGTGCA CTTACCCCTTC TGTTCCTCTT TAAGGACCTT  
  
GTCTTCCCTT GCAGGCTACA GCAAGTGCAG CCGCTGCAGG AGCGTCTCTT 150  
CAGAACGGCA CGTGGGTGT CATTACGTTC GCGGACGTCC TCGCAGGAGA  
  
TGCCACTGCC CGGCTCAGCC ACCACCAGCA CAGTGTCCGC ATCGTCATTG 200  
ACGGTCAAGG GCGGAGTGGG TGGTGGTGGT GTACACAGCG TAGCAGTAAC  
  
AGGTGGCCAG GGGTGCCTGAA CAGCTCTCTC AGGCCCCAGG TGGCTGGGCT 250  
TCCACCGGTC CCGACGACTT CTCGAGGAGG TCGGGGTCCC ACCGACCCGA  
  
CTTCTGCGGg ggtccagcca tgcacacatc TCGCCAGACC TCCAGGACAT 300  
GAAGACGGC ccaggtcggc acgggtgtag ACCGGTCTGG AGGTCTGTGA  
  
TCTCTGTGTA TAATGCTCTC AGGCAGAGCG TCTCTGCTCC ATCATAGGTA 350  
AGACACACAT ATACAGGAGG TCCGTCTCGC AGAGACGAGG TAGTATCCAT  
  
CTGAGGAAGC GAGACTGAGC AGACACCGTG GTCTTCAGCT TGGCCATATA 400  
GACTCCTTCG CTGACTCG TCTGTGGCAC CAGGAGTGA ACCGGTATAT  
  
CTTCTTGCA TGGGCAGCTG GAAGGCAGAA GAAGAGGAG ATGAAGGTTG 450  
GAACAACGTA CACCGTCGAC CTTCGGTCTT CTCTCCGTTC TACTTCACC  
  
CACCATGGTG AAGACGGGAC CTAACCAGAC AATGGGCTCC TCGGGGGGAC 500  
GTGGTACCAC TTCTGCCCCG GATTGGTCTG TTACCCGAGG ACCGCCCCTG  
  
GCTGACATAA CTGAAGGCAT AGGACAGCCA GCGGGGGGGG3'-SEQ ID NO 7 540  
CGACTGTATT GACTTCGCTA TCTCTCGGT CCGCGGGGGG5'-SEQ ID NO 8

9/15

FIGURE 9

JW1

```
5' GCACTGGGCA CCCACTACCA ATGGATTGGA ATTGGTCCTT AAGATAAAAT 50
3' CGTGACCCGT GGGTGATGGT TACCTAACCT TAACCAGGAA TTCTATTTTA

GTACCTGATC CAGCCCAATA TCTTCAATTT ACAGATACTG TATCAAAACC 100
CATGGACTAG GTCCGGTTAT AGAAGTTAAA TGTCTATGAC ATAGTTTIGG

CTGAGCGAC AAGGGACATT TCCAAGTAC CCAGAaagac togagtgtcc 150
GACTCTOCTG TTCCCTGTAA AGGTTTCAGTG GGTCtttctg aggtcacagy

TCTCTTGAAA TCCAATGGTC TTTTTCCTT ACTCCATTGC CTAACATTGT 200
aCAGAACTTT AGGTTACCAG AAAAAAGGAA TGAGGTAAAG GATTGTAAACA

GGGGTAGAAA TAAAGTTCAA AGACCTTCAG AACTGGCCCC AGCTOCTCCC 250
CCCCATCTTT ATTTCAAGTT TCTGGAAGTC TTGACCGGGG TOGAGCAGGG

TCTTCAOCTG ATCTCCCCAA GAAACTGCA GGATAGACTC TGAAGCTTAC 300
AGAAGTGGAC TAGAGGGGTT CTTTTCAGCT CCTATCTGAG ACTTCCAATG

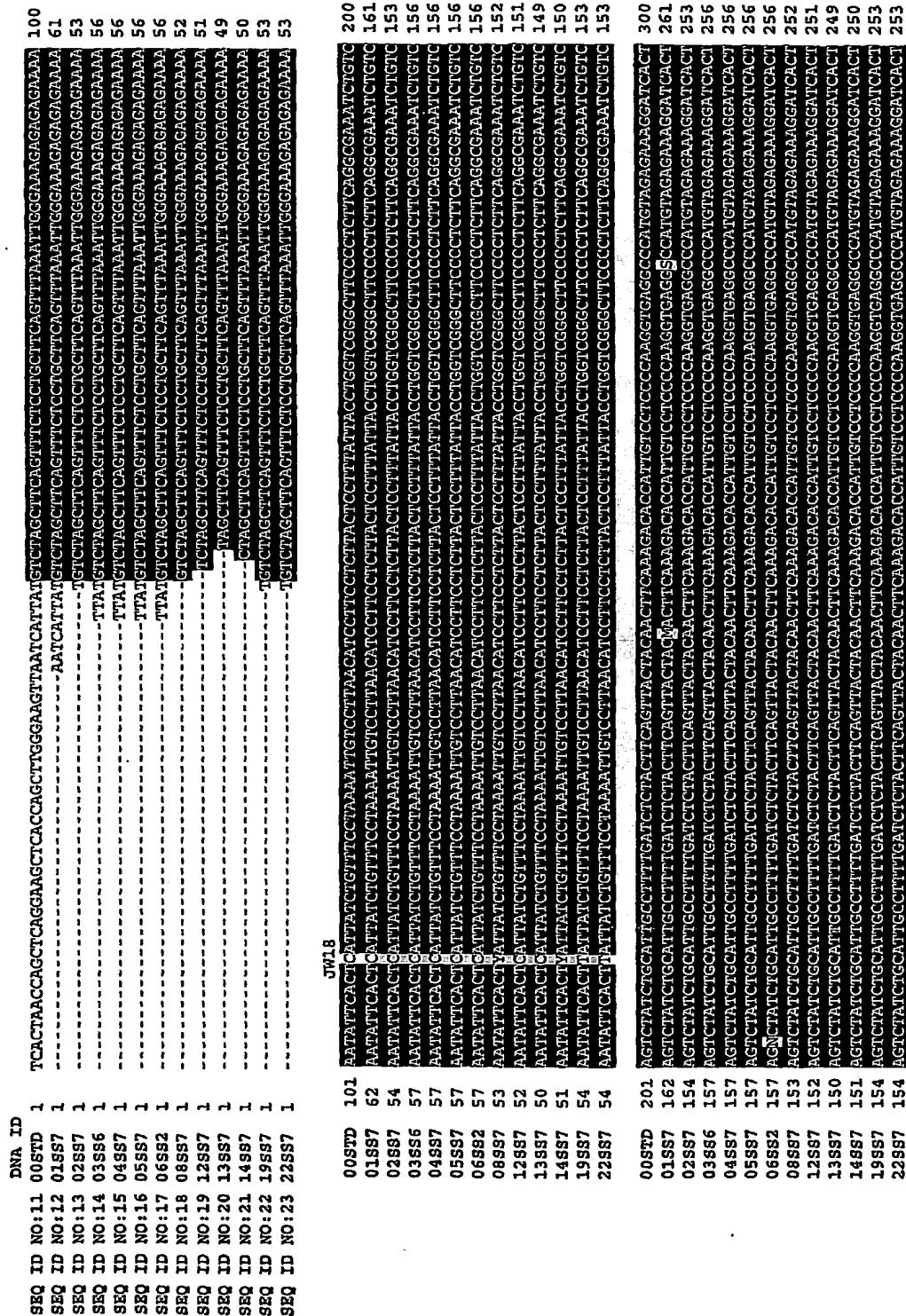
CTGAGCCACC TCAAGCTCTG GIGATCACCC AAGGCTTCAG CCAGGGCOCTG 350
GACTCGGTGG AGTTGAGAC CACTAGTGGG TTCCGAAGTC GGTCOCGGAC

GGCCCCCTOG TCACCCACTC TGTGCCCCA GAATCTGAAA AGGCCAAAAG 400
CCGGGGGAGC AGTGGGTGAG ACAACGGGGT CTTAGACTTT TCGGTTTTTC

AGTCAACAGA CAGTGTCACT GAGTACCTGA TATGTGTCT AGACATGAAC 450
TCAGTTGTCT GTCACAGTCA CTCATGGACT ATACACAAGA TCTGTACTTG

TAACAGTCCT CCTCCCTCTG CAGTCCCAGC CAGAGGGGCA GGACCACTCA 500
ATTGTACAGG GGAGGGAGAC GTCAGGGTGG GTCTCCCGGT CTTGGTCACT

ATCCCAAGT GGCCTCACTG 3'-SEQ ID NO 9 520
TAGGGTCTCA CCGGAGTAC 5'-SEQ ID NO 10
```



**FIGURE 10**

Inventors: Abreu et al.  
 Filing Date: Herewith  
 Attorney Docket: P-CE 5480  
 USPTO Customer No. 23601 (858-535-9001)

11/15

0087D 301 TCCTTGCTGAAAGAGAGGTCACAGGGGCAACCCACGTGGGGCCCTCCCTGAAACCCAGGCCCCAGCCCTGAGCCTGAGACACCTCTTCTTCTTCTGAGAGCCAC 400  
 01587 262 TCCTTGCTGAAAGAGAGGTCACAGGGGCAACCCACGTGGGGCCCTCCCTGAAACCCAGGCCCCAGCCCTGAGCCTGAGACACCTCTTCTTCTTCTGAGAGCCAC 361  
 02887 254 TCCTTGCTGAAAGAGAGGTCACAGGGGCAACCCACGTGGGGCCCTCCCTGAAACCCAGGCCCCAGCCCTGAGCCTGAGACACCTCTTCTTCTTCTGAGAGCCAC 353  
 03886 257 TCCTTGCTGAAAGAGAGGTCACAGGGGCAACCCACGTGGGGCCCTCCCTGAAACCCAGGCCCCAGCCCTGAGCCTGAGACACCTCTTCTTCTTCTGAGAGCCAC 356  
 04887 257 TCCTTGCTGAAAGAGAGGTCACAGGGGCAACCCACGTGGGGCCCTCCCTGAAACCCAGGCCCCAGCCCTGAGCCTGAGACACCTCTTCTTCTTCTGAGAGCCAC 356  
 05887 257 TCCTTGCTGAAAGAGAGGTCACAGGGGCAACCCACGTGGGGCCCTCCCTGAAACCCAGGCCCCAGCCCTGAGCCTGAGACACCTCTTCTTCTTCTGAGAGCCAC 356  
 06882 257 TCCTTGCTGAAAGAGAGGTCACAGGGGCAACCCACGTGGGGCCCTCCCTGAAACCCAGGCCCCAGCCCTGAGCCTGAGACACCTCTTCTTCTTCTGAGAGCCAC 356  
 08887 252 TCCTTGCTGAAAGAGAGGTCACAGGGGCAACCCACGTGGGGCCCTCCCTGAAACCCAGGCCCCAGCCCTGAGCCTGAGACACCTCTTCTTCTTCTGAGAGCCAC 352  
 12887 252 TCCTTGCTGAAAGAGAGGTCACAGGGGCAACCCACGTGGGGCCCTCCCTGAAACCCAGGCCCCAGCCCTGAGCCTGAGACACCTCTTCTTCTTCTGAGAGCCAC 351  
 13887 250 TCCTTGCTGAAAGAGAGGTCACAGGGGCAACCCACGTGGGGCCCTCCCTGAAACCCAGGCCCCAGCCCTGAGCCTGAGACACCTCTTCTTCTTCTGAGAGCCAC 349  
 14887 251 TCCTTGCTGAAAGAGAGGTCACAGGGGCAACCCACGTGGGGCCCTCCCTGAAACCCAGGCCCCAGCCCTGAGCCTGAGACACCTCTTCTTCTTCTGAGAGCCAC 350  
 19887 254 TCCTTGCTGAAAGAGAGGTCACAGGGGCAACCCACGTGGGGCCCTCCCTGAAACCCAGGCCCCAGCCCTGAGCCTGAGACACCTCTTCTTCTTCTGAGAGCCAC 353  
 22887 254 TCCTTGCTGAAAGAGAGGTCACAGGGGCAACCCACGTGGGGCCCTCCCTGAAACCCAGGCCCCAGCCCTGAGCCTGAGACACCTCTTCTTCTTCTGAGAGCCAC 353

JWL7

0087D 401 AGCCAGCCCCGGTTCTCTGGGGCCCAAGAGCAATGCTTTGCTTAAGCTGTAATCTCAGCCCACTGACCCCTTTCGAGACAGGAGAGGGGGAAGCCCA 500  
 01887 362 AGCCAGCCCCGGTTCTCTGGGGCCCAAGAGCAATGCTTTGCTTAAGCTGTAATCTCAGCCCACTGACCCCTTTCGAGACAGGAGAGGGGGAAGCCCA 461  
 02887 354 AGCCAGCCCCGGTTCTCTGGGGCCCAAGAGCAATGCTTTGCTTAAGCTGTAATCTCAGCCCACTGACCCCTTTCGAGACAGGAGAGGGGGAAGCCCA 453  
 03886 357 AGCCAGCCCCGGTTCTCTGGGGCCCAAGAGCAATGCTTTGCTTAAGCTGTAATCTCAGCCCACTGACCCCTTTCGAGACAGGAGAGGGGGAAGCCCA 456  
 04887 357 AGCCAGCCCCGGTTCTCTGGGGCCCAAGAGCAATGCTTTGCTTAAGCTGTAATCTCAGCCCACTGACCCCTTTCGAGACAGGAGAGGGGGAAGCCCA 456  
 05887 357 AGCCAGCCCCGGTTCTCTGGGGCCCAAGAGCAATGCTTTGCTTAAGCTGTAATCTCAGCCCACTGACCCCTTTCGAGACAGGAGAGGGGGAAGCCCA 456  
 06882 357 AGCCAGCCCCGGTTCTCTGGGGCCCAAGAGCAATGCTTTGCTTAAGCTGTAATCTCAGCCCACTGACCCCTTTCGAGACAGGAGAGGGGGAAGCCCA 456  
 08887 353 AGCCAGCCCCGGTTCTCTGGGGCCCAAGAGCAATGCTTTGCTTAAGCTGTAATCTCAGCCCACTGACCCCTTTCGAGACAGGAGAGGGGGAAGCCCA 452  
 12887 352 AGCCAGCCCCGGTTCTCTGGGGCCCAAGAGCAATGCTTTGCTTAAGCTGTAATCTCAGCCCACTGACCCCTTTCGAGACAGGAGAGGGGGAAGCCCA 451  
 13887 350 AGCCAGCCCCGGTTCTCTGGGGCCCAAGAGCAATGCTTTGCTTAAGCTGTAATCTCAGCCCACTGACCCCTTTCGAGACAGGAGAGGGGGAAGCCCA 449  
 14887 351 AGCCAGCCCCGGTTCTCTGGGGCCCAAGAGCAATGCTTTGCTTAAGCTGTAATCTCAGCCCACTGACCCCTTTCGAGACAGGAGAGGGGGAAGCCCA 450  
 19887 354 AGCCAGCCCCGGTTCTCTGGGGCCCAAGAGCAATGCTTTGCTTAAGCTGTAATCTCAGCCCACTGACCCCTTTCGAGACAGGAGAGGGGGAAGCCCA 453  
 22887 354 AGCCAGCCCCGGTTCTCTGGGGCCCAAGAGCAATGCTTTGCTTAAGCTGTAATCTCAGCCCACTGACCCCTTTCGAGACAGGAGAGGGGGAAGCCCA 453

0087D 501 GGGAGGCTCAACTTCCCAAGTGTCTGAGTCTCTG 535  
 01887 462 GGGAGGCTCAACTTCCCAAGTGTCTGAGTCTCTG 496  
 02887 454 GGGAGGCTCAACTTCCCAAGTGTCTGAGTCTCTG 488  
 03886 457 GGGAGGCTCAACTTCCCAAGTGTCTGAGTCTCTG 491  
 04887 457 GGGAGGCTCAACTTCCCAAGTGTCTGAGTCTCTG 491  
 05887 457 GGGAGGCTCAACTTCCCAAGTGTCTGAGTCTCTG 491  
 06882 457 GGGAGGCTCAACTTCCCAAGTGTCTGAGTCTCTG 491  
 08887 453 GGGAGGCTCAACTTCCCAAGTGTCTGAGTCTCTG 487  
 12887 452 GGGAGGCTCAACTTCCCAAGTGTCTGAGTCTCTG 486  
 13887 450 GGGAGGCTCAACTTCCCAAGTGTCTGAGTCTCTG 484  
 14887 451 GGGAGGCTCAACTTCCCAAGTGTCTGAGTCTCTG 485  
 19887 454 GGGAGGCTCAACTTCCCAAGTGTCTGAGTCTCTG 488  
 22887 454 GGGAGGCTCAACTTCCCAAGTGTCTGAGTCTCTG 488

FIGURE 10 (continued)

12/15

DNA ID	100	TCACAGACAGCTCTGAAATAGACCCACAGAGGAGGAGTCCATTTCAATTTGTCGCCAATGCT	100
SEQ ID NO:24	008TD	TCACAGACAGCTCTGAAATAGACCCACAGAGGAGGAGTCCATTTCAATTTGTCGCCAATGCT	63
SEQ ID NO:25	018S6	TCACAGACAGCTCTGAAATAGACCCACAGAGGAGGAGTCCATTTCAATTTGTCGCCAATGCT	65
SEQ ID NO:26	028S6	TCACAGACAGCTCTGAAATAGACCCACAGAGGAGGAGTCCATTTCAATTTGTCGCCAATGCT	62
SEQ ID NO:27	038S6	TCACAGACAGCTCTGAAATAGACCCACAGAGGAGGAGTCCATTTCAATTTGTCGCCAATGCT	70
SEQ ID NO:28	048S6	TCACAGACAGCTCTGAAATAGACCCACAGAGGAGGAGTCCATTTCAATTTGTCGCCAATGCT	70
SEQ ID NO:29	058S2	TCACAGACAGCTCTGAAATAGACCCACAGAGGAGGAGTCCATTTCAATTTGTCGCCAATGCT	69
SEQ ID NO:30	068S6	TCACAGACAGCTCTGAAATAGACCCACAGAGGAGGAGTCCATTTCAATTTGTCGCCAATGCT	69
SEQ ID NO:31	088S6	TCACAGACAGCTCTGAAATAGACCCACAGAGGAGGAGTCCATTTCAATTTGTCGCCAATGCT	63
SEQ ID NO:32	128S6	TCACAGACAGCTCTGAAATAGACCCACAGAGGAGGAGTCCATTTCAATTTGTCGCCAATGCT	70
SEQ ID NO:33	138S6	TCACAGACAGCTCTGAAATAGACCCACAGAGGAGGAGTCCATTTCAATTTGTCGCCAATGCT	63
SEQ ID NO:34	148S6	TCACAGACAGCTCTGAAATAGACCCACAGAGGAGGAGTCCATTTCAATTTGTCGCCAATGCT	65
SEQ ID NO:35	198S6	TCACAGACAGCTCTGAAATAGACCCACAGAGGAGGAGTCCATTTCAATTTGTCGCCAATGCT	66
SEQ ID NO:36	228S6	TCACAGACAGCTCTGAAATAGACCCACAGAGGAGGAGTCCATTTCAATTTGTCGCCAATGCT	66
DNA ID	101	TTAGGATGTACAGTATGGATTGAAAGTTTACAGGAAATTAAGGCGGTTCCTTCAAGCAAAATGCTCTCTGGATTATTTCAAAATGATGATGTTG	200
SEQ ID NO:24	008TD	TTAGGATGTACAGTATGGATTGAAAGTTTACAGGAAATTAAGGCGGTTCCTTCAAGCAAAATGCTCTCTGGATTATTTCAAAATGATGATGTTG	163
SEQ ID NO:25	018S6	TTAGGATGTACAGTATGGATTGAAAGTTTACAGGAAATTAAGGCGGTTCCTTCAAGCAAAATGCTCTCTGGATTATTTCAAAATGATGATGTTG	165
SEQ ID NO:26	028S6	TTAGGATGTACAGTATGGATTGAAAGTTTACAGGAAATTAAGGCGGTTCCTTCAAGCAAAATGCTCTCTGGATTATTTCAAAATGATGATGTTG	162
SEQ ID NO:27	038S6	TTAGGATGTACAGTATGGATTGAAAGTTTACAGGAAATTAAGGCGGTTCCTTCAAGCAAAATGCTCTCTGGATTATTTCAAAATGATGATGTTG	170
SEQ ID NO:28	048S6	TTAGGATGTACAGTATGGATTGAAAGTTTACAGGAAATTAAGGCGGTTCCTTCAAGCAAAATGCTCTCTGGATTATTTCAAAATGATGATGTTG	170
SEQ ID NO:29	058S2	TTAGGATGTACAGTATGGATTGAAAGTTTACAGGAAATTAAGGCGGTTCCTTCAAGCAAAATGCTCTCTGGATTATTTCAAAATGATGATGTTG	169
SEQ ID NO:30	068S6	TTAGGATGTACAGTATGGATTGAAAGTTTACAGGAAATTAAGGCGGTTCCTTCAAGCAAAATGCTCTCTGGATTATTTCAAAATGATGATGTTG	163
SEQ ID NO:31	088S6	TTAGGATGTACAGTATGGATTGAAAGTTTACAGGAAATTAAGGCGGTTCCTTCAAGCAAAATGCTCTCTGGATTATTTCAAAATGATGATGTTG	170
SEQ ID NO:32	128S6	TTAGGATGTACAGTATGGATTGAAAGTTTACAGGAAATTAAGGCGGTTCCTTCAAGCAAAATGCTCTCTGGATTATTTCAAAATGATGATGTTG	163
SEQ ID NO:33	138S6	TTAGGATGTACAGTATGGATTGAAAGTTTACAGGAAATTAAGGCGGTTCCTTCAAGCAAAATGCTCTCTGGATTATTTCAAAATGATGATGTTG	163
SEQ ID NO:34	148S6	TTAGGATGTACAGTATGGATTGAAAGTTTACAGGAAATTAAGGCGGTTCCTTCAAGCAAAATGCTCTCTGGATTATTTCAAAATGATGATGTTG	165
SEQ ID NO:35	198S6	TTAGGATGTACAGTATGGATTGAAAGTTTACAGGAAATTAAGGCGGTTCCTTCAAGCAAAATGCTCTCTGGATTATTTCAAAATGATGATGTTG	165
SEQ ID NO:36	228S6	TTAGGATGTACAGTATGGATTGAAAGTTTACAGGAAATTAAGGCGGTTCCTTCAAGCAAAATGCTCTCTGGATTATTTCAAAATGATGATGTTG	166
DNA ID	201	AAGCCTTTGTAATTTGTACAGATGCTGTGCAAAATGTTATTTTAAACATTTATGATGTGTGAAAACCTGGTTAATATTTATAGTTCACCTTTGTTTACTGT	300
SEQ ID NO:24	008TD	AAGCCTTTGTAATTTGTACAGATGCTGTGCAAAATGTTATTTTAAACATTTATGATGTGTGAAAACCTGGTTAATATTTATAGTTCACCTTTGTTTACTGT	263
SEQ ID NO:25	018S6	AAGCCTTTGTAATTTGTACAGATGCTGTGCAAAATGTTATTTTAAACATTTATGATGTGTGAAAACCTGGTTAATATTTATAGTTCACCTTTGTTTACTGT	265
SEQ ID NO:26	028S6	AAGCCTTTGTAATTTGTACAGATGCTGTGCAAAATGTTATTTTAAACATTTATGATGTGTGAAAACCTGGTTAATATTTATAGTTCACCTTTGTTTACTGT	262
SEQ ID NO:27	038S6	AAGCCTTTGTAATTTGTACAGATGCTGTGCAAAATGTTATTTTAAACATTTATGATGTGTGAAAACCTGGTTAATATTTATAGTTCACCTTTGTTTACTGT	270
SEQ ID NO:28	048S6	AAGCCTTTGTAATTTGTACAGATGCTGTGCAAAATGTTATTTTAAACATTTATGATGTGTGAAAACCTGGTTAATATTTATAGTTCACCTTTGTTTACTGT	270
SEQ ID NO:29	058S2	AAGCCTTTGTAATTTGTACAGATGCTGTGCAAAATGTTATTTTAAACATTTATGATGTGTGAAAACCTGGTTAATATTTATAGTTCACCTTTGTTTACTGT	269
SEQ ID NO:30	068S6	AAGCCTTTGTAATTTGTACAGATGCTGTGCAAAATGTTATTTTAAACATTTATGATGTGTGAAAACCTGGTTAATATTTATAGTTCACCTTTGTTTACTGT	269
SEQ ID NO:31	088S6	AAGCCTTTGTAATTTGTACAGATGCTGTGCAAAATGTTATTTTAAACATTTATGATGTGTGAAAACCTGGTTAATATTTATAGTTCACCTTTGTTTACTGT	263
SEQ ID NO:32	128S6	AAGCCTTTGTAATTTGTACAGATGCTGTGCAAAATGTTATTTTAAACATTTATGATGTGTGAAAACCTGGTTAATATTTATAGTTCACCTTTGTTTACTGT	270
SEQ ID NO:33	138S6	AAGCCTTTGTAATTTGTACAGATGCTGTGCAAAATGTTATTTTAAACATTTATGATGTGTGAAAACCTGGTTAATATTTATAGTTCACCTTTGTTTACTGT	263
SEQ ID NO:34	148S6	AAGCCTTTGTAATTTGTACAGATGCTGTGCAAAATGTTATTTTAAACATTTATGATGTGTGAAAACCTGGTTAATATTTATAGTTCACCTTTGTTTACTGT	265
SEQ ID NO:35	198S6	AAGCCTTTGTAATTTGTACAGATGCTGTGCAAAATGTTATTTTAAACATTTATGATGTGTGAAAACCTGGTTAATATTTATAGTTCACCTTTGTTTACTGT	266
SEQ ID NO:36	228S6	AAGCCTTTGTAATTTGTACAGATGCTGTGCAAAATGTTATTTTAAACATTTATGATGTGTGAAAACCTGGTTAATATTTATAGTTCACCTTTGTTTACTGT	266

FIGURE 11A

13/15

DNA ID	301	CTTAAGTTTATCTCTATAGCAACATGCGCCGTTGAACTTTATCTGTAATAATATCAGAGGGGAATAAATCTGTGAGTCAAAAACAGCCATCTTCTTGTG	400
09STD	264	CTTAAGTTTATCTCTATAGCAACATGCGCCGTTGAACTTTATCTGTAATAATATCAGAGGGGAATAAATCTGTGAGTCAAAAACAGCCATCTTCTTGTG	363
01SS6	266	CTTAAGTTTATCTCTATAGCAACATGCGCCGTTGAACTTTATCTGTAATAATATCAGAGGGGAATAAATCTGTGAGTCAAAAACAGCCATCTTCTTGTG	365
02SS6	266	CTTAAGTTTATCTCTATAGCAACATGCGCCGTTGAACTTTATCTGTAATAATATCAGAGGGGAATAAATCTGTGAGTCAAAAACAGCCATCTTCTTGTG	365
03SS6	263	CTTAAGTTTATCTCTATAGCAACATGCGCCGTTGAACTTTATCTGTAATAATATCAGAGGGGAATAAATCTGTGAGTCAAAAACAGCCATCTTCTTGTG	362
04SS6	271	CTTAAGTTTATCTCTATAGCAACATGCGCCGTTGAACTTTATCTGTAATAATATCAGAGGGGAATAAATCTGTGAGTCAAAAACAGCCATCTTCTTGTG	370
05SS2	271	CTTAAGTTTATCTCTATAGCAACATGCGCCGTTGAACTTTATCTGTAATAATATCAGAGGGGAATAAATCTGTGAGTCAAAAACAGCCATCTTCTTGTG	369
06SS6	270	CTTAAGTTTATCTCTATAGCAACATGCGCCGTTGAACTTTATCTGTAATAATATCAGAGGGGAATAAATCTGTGAGTCAAAAACAGCCATCTTCTTGTG	369
08SS6	270	CTTAAGTTTATCTCTATAGCAACATGCGCCGTTGAACTTTATCTGTAATAATATCAGAGGGGAATAAATCTGTGAGTCAAAAACAGCCATCTTCTTGTG	369
12SS6	264	CTTAAGTTTATCTCTATAGCAACATGCGCCGTTGAACTTTATCTGTAATAATATCAGAGGGGAATAAATCTGTGAGTCAAAAACAGCCATCTTCTTGTG	363
13SS6	271	CTTAAGTTTATCTCTATAGCAACATGCGCCGTTGAACTTTATCTGTAATAATATCAGAGGGGAATAAATCTGTGAGTCAAAAACAGCCATCTTCTTGTG	370
14SS6	264	CTTAAGTTTATCTCTATAGCAACATGCGCCGTTGAACTTTATCTGTAATAATATCAGAGGGGAATAAATCTGTGAGTCAAAAACAGCCATCTTCTTGTG	363
19SS6	266	CTTAAGTTTATCTCTATAGCAACATGCGCCGTTGAACTTTATCTGTAATAATATCAGAGGGGAATAAATCTGTGAGTCAAAAACAGCCATCTTCTTGTG	365
22SS6	267	CTTAAGTTTATCTCTATAGCAACATGCGCCGTTGAACTTTATCTGTAATAATATCAGAGGGGAATAAATCTGTGAGTCAAAAACAGCCATCTTCTTGTG	366
00STD	401	ACCAACATTTTAAATAATATCTGCTGGGCAACATGCTGCTCAGGCTGTAATCCACGACATTTGGGAGGCGGAGTGGGACATCACCTGAGGTTGGG	497
01SS6	364	ACCAACATTTTAAATAATATCTGCTGGGCAACATGCTGCTCAGGCTGTAATCCACGACATTTGGGAGGCGGAGTGGGACATCACCTGAGGTTGGG	460
02SS6	366	ACCAACATTTTAAATAATATCTGCTGGGCAACATGCTGCTCAGGCTGTAATCCACGACATTTGGGAGGCGGAGTGGGACATCACCTGAGGTTGGG	462
03SS6	363	ACCAACATTTTAAATAATATCTGCTGGGCAACATGCTGCTCAGGCTGTAATCCACGACATTTGGGAGGCGGAGTGGGACATCACCTGAGGTTGGG	459
04SS6	371	ACCAACATTTTAAATAATATCTGCTGGGCAACATGCTGCTCAGGCTGTAATCCACGACATTTGGGAGGCGGAGTGGGACATCACCTGAGGTTGGG	467
05SS2	371	ACCAACATTTTAAATAATATCTGCTGGGCAACATGCTGCTCAGGCTGTAATCCACGACATTTGGGAGGCGGAGTGGGACATCACCTGAGGTTGGG	467
06SS6	370	ACCAACATTTTAAATAATATCTGCTGGGCAACATGCTGCTCAGGCTGTAATCCACGACATTTGGGAGGCGGAGTGGGACATCACCTGAGGTTGGG	466
08SS6	370	ACCAACATTTTAAATAATATCTGCTGGGCAACATGCTGCTCAGGCTGTAATCCACGACATTTGGGAGGCGGAGTGGGACATCACCTGAGGTTGGG	466
12SS6	364	ACCAACATTTTAAATAATATCTGCTGGGCAACATGCTGCTCAGGCTGTAATCCACGACATTTGGGAGGCGGAGTGGGACATCACCTGAGGTTGGG	460
13SS6	371	ACCAACATTTTAAATAATATCTGCTGGGCAACATGCTGCTCAGGCTGTAATCCACGACATTTGGGAGGCGGAGTGGGACATCACCTGAGGTTGGG	467
14SS6	364	ACCAACATTTTAAATAATATCTGCTGGGCAACATGCTGCTCAGGCTGTAATCCACGACATTTGGGAGGCGGAGTGGGACATCACCTGAGGTTGGG	460
19SS6	366	ACCAACATTTTAAATAATATCTGCTGGGCAACATGCTGCTCAGGCTGTAATCCACGACATTTGGGAGGCGGAGTGGGACATCACCTGAGGTTGGG	462
22SS6	367	ACCAACATTTTAAATAATATCTGCTGGGCAACATGCTGCTCAGGCTGTAATCCACGACATTTGGGAGGCGGAGTGGGACATCACCTGAGGTTGGG	463

FIGURE 11A (continued)

14/15

SEQ ID NO:55	009TD	1	TTGGTCTCAGTTTGTGAGCAGGCTGTGAGTTT	100
SEQ ID NO:56	019S6	1	GGGCCCCAGAGGCTGGGTGACATGTTGGCAGGCTCTTCAAAATGAGCCCTGCTGCTTAAG	64
SEQ ID NO:57	025S6	1	GGGCCCCAGAGGCTGGGTGACATGTTGGCAGGCTCTTCAAAATGAGCCCTGCTGCTTAAG	64
SEQ ID NO:58	038S6	1	GGGCCCCAGAGGCTGGGTGACATGTTGGCAGGCTCTTCAAAATGAGCCCTGCTGCTTAAG	64
SEQ ID NO:59	045S6	1	GGGCCCCAGAGGCTGGGTGACATGTTGGCAGGCTCTTCAAAATGAGCCCTGCTGCTTAAG	50
SEQ ID NO:60	058S2	1	GGGCCCCAGAGGCTGGGTGACATGTTGGCAGGCTCTTCAAAATGAGCCCTGCTGCTTAAG	67
SEQ ID NO:61	065S6	1	GGGCCCCAGAGGCTGGGTGACATGTTGGCAGGCTCTTCAAAATGAGCCCTGCTGCTTAAG	66
SEQ ID NO:62	088S6	1	GGGCCCCAGAGGCTGGGTGACATGTTGGCAGGCTCTTCAAAATGAGCCCTGCTGCTTAAG	65
SEQ ID NO:63	128S6	1	GGGCCCCAGAGGCTGGGTGACATGTTGGCAGGCTCTTCAAAATGAGCCCTGCTGCTTAAG	65
SEQ ID NO:64	138S6	1	GGGCCCCAGAGGCTGGGTGACATGTTGGCAGGCTCTTCAAAATGAGCCCTGCTGCTTAAG	70
SEQ ID NO:65	148S6	1	GGGCCCCAGAGGCTGGGTGACATGTTGGCAGGCTCTTCAAAATGAGCCCTGCTGCTTAAG	20
SEQ ID NO:66	198S6	1	GGGCCCCAGAGGCTGGGTGACATGTTGGCAGGCTCTTCAAAATGAGCCCTGCTGCTTAAG	67
SEQ ID NO:67	228S6	1	GGGCCCCAGAGGCTGGGTGACATGTTGGCAGGCTCTTCAAAATGAGCCCTGCTGCTTAAG	67
009TD	101		GCTGAACCTTGTCTTGGGAACACCAATAGTCACTTATTTCTGGCAGAGGAGGAGCATCAGTGCCTCCAGGATAGACTTTTCCCAAGCCTACTTTTG	200
019S6	65		GCTGAACCTTGTCTTGGGAACACCAATAGTCACTTATTTCTGGCAGAGGAGGAGCATCAGTGCCTCCAGGATAGACTTTTCCCAAGCCTACTTTTG	164
028S6	65		GCTGAACCTTGTCTTGGGAACACCAATAGTCACTTATTTCTGGCAGAGGAGGAGCATCAGTGCCTCCAGGATAGACTTTTCCCAAGCCTACTTTTG	164
038S6	65		GCTGAACCTTGTCTTGGGAACACCAATAGTCACTTATTTCTGGCAGAGGAGGAGCATCAGTGCCTCCAGGATAGACTTTTCCCAAGCCTACTTTTG	164
048S6	51		GCTGAACCTTGTCTTGGGAACACCAATAGTCACTTATTTCTGGCAGAGGAGGAGCATCAGTGCCTCCAGGATAGACTTTTCCCAAGCCTACTTTTG	150
058S2	68		GCTGAACCTTGTCTTGGGAACACCAATAGTCACTTATTTCTGGCAGAGGAGGAGCATCAGTGCCTCCAGGATAGACTTTTCCCAAGCCTACTTTTG	166
068S6	67		GCTGAACCTTGTCTTGGGAACACCAATAGTCACTTATTTCTGGCAGAGGAGGAGCATCAGTGCCTCCAGGATAGACTTTTCCCAAGCCTACTTTTG	165
088S6	66		GCTGAACCTTGTCTTGGGAACACCAATAGTCACTTATTTCTGGCAGAGGAGGAGCATCAGTGCCTCCAGGATAGACTTTTCCCAAGCCTACTTTTG	165
128S6	66		GCTGAACCTTGTCTTGGGAACACCAATAGTCACTTATTTCTGGCAGAGGAGGAGCATCAGTGCCTCCAGGATAGACTTTTCCCAAGCCTACTTTTG	165
138S6	66		GCTGAACCTTGTCTTGGGAACACCAATAGTCACTTATTTCTGGCAGAGGAGGAGCATCAGTGCCTCCAGGATAGACTTTTCCCAAGCCTACTTTTG	170
148S6	71		GCTGAACCTTGTCTTGGGAACACCAATAGTCACTTATTTCTGGCAGAGGAGGAGCATCAGTGCCTCCAGGATAGACTTTTCCCAAGCCTACTTTTG	120
198S6	21		GCTGAACCTTGTCTTGGGAACACCAATAGTCACTTATTTCTGGCAGAGGAGGAGCATCAGTGCCTCCAGGATAGACTTTTCCCAAGCCTACTTTTG	167
228S6	68		GCTGAACCTTGTCTTGGGAACACCAATAGTCACTTATTTCTGGCAGAGGAGGAGCATCAGTGCCTCCAGGATAGACTTTTCCCAAGCCTACTTTTG	167
009TD	201		CCATTGACTTCTTCCCAAGATTCATATCCAGGATGTACAAGGACAGCCCTCTCTCCATAGTAGGAGCTGGCTCTGCTGATCTCTCCAGGCTTCGGTGT	300
019S6	165		CCATTGACTTCTTCCCAAGATTCATATCCAGGATGTACAAGGACAGCCCTCTCTCCATAGTAGGAGCTGGCTCTGCTGATCTCTCCAGGCTTCGGTGT	264
028S6	165		CCATTGACTTCTTCCCAAGATTCATATCCAGGATGTACAAGGACAGCCCTCTCTCCATAGTAGGAGCTGGCTCTGCTGATCTCTCCAGGCTTCGGTGT	264
038S6	165		CCATTGACTTCTTCCCAAGATTCATATCCAGGATGTACAAGGACAGCCCTCTCTCCATAGTAGGAGCTGGCTCTGCTGATCTCTCCAGGCTTCGGTGT	264
048S6	151		CCATTGACTTCTTCCCAAGATTCATATCCAGGATGTACAAGGACAGCCCTCTCTCCATAGTAGGAGCTGGCTCTGCTGATCTCTCCAGGCTTCGGTGT	250
058S2	168		CCATTGACTTCTTCCCAAGATTCATATCCAGGATGTACAAGGACAGCCCTCTCTCCATAGTAGGAGCTGGCTCTGCTGATCTCTCCAGGCTTCGGTGT	267
068S6	167		CCATTGACTTCTTCCCAAGATTCATATCCAGGATGTACAAGGACAGCCCTCTCTCCATAGTAGGAGCTGGCTCTGCTGATCTCTCCAGGCTTCGGTGT	266
088S6	166		CCATTGACTTCTTCCCAAGATTCATATCCAGGATGTACAAGGACAGCCCTCTCTCCATAGTAGGAGCTGGCTCTGCTGATCTCTCCAGGCTTCGGTGT	265
128S6	166		CCATTGACTTCTTCCCAAGATTCATATCCAGGATGTACAAGGACAGCCCTCTCTCCATAGTAGGAGCTGGCTCTGCTGATCTCTCCAGGCTTCGGTGT	265
138S6	166		CCATTGACTTCTTCCCAAGATTCATATCCAGGATGTACAAGGACAGCCCTCTCTCCATAGTAGGAGCTGGCTCTGCTGATCTCTCCAGGCTTCGGTGT	265
148S6	171		CCATTGACTTCTTCCCAAGATTCATATCCAGGATGTACAAGGACAGCCCTCTCTCCATAGTAGGAGCTGGCTCTGCTGATCTCTCCAGGCTTCGGTGT	270
198S6	121		CCATTGACTTCTTCCCAAGATTCATATCCAGGATGTACAAGGACAGCCCTCTCTCCATAGTAGGAGCTGGCTCTGCTGATCTCTCCAGGCTTCGGTGT	220
228S6	168		CCATTGACTTCTTCCCAAGATTCATATCCAGGATGTACAAGGACAGCCCTCTCTCCATAGTAGGAGCTGGCTCTGCTGATCTCTCCAGGCTTCGGTGT	257

FIGURE 11B



Inventors: Abreu et al.

**Filing Date: Herewith**

Attorney Docket: P-CE 5480

USPTO Customer No. 23601 (858-535-9001)

15/15

JW15

DNA ID	301	GGGTCACTGGGGCCCATGGATGTCGTTGCTTAACCTGAGTGCCTTTTGGTGGAGAGGCGCGGCTCTCTCAAAAGAGACCCCTTACCACTGGCTCTGATGAAGAG	400
008TD	301	GGGTCACTGGGGCCCATGGATGTCGTTGCTTAACCTGAGTGCCTTTTGGTGGAGAGGCGCGGCTCTCTCAAAAGAGACCCCTTACCACTGGCTCTGATGAAGAG	364
01886	265	GGGTCACTGGGGCCCATGGATGTCGTTGCTTAACCTGAGTGCCTTTTGGTGGAGAGGCGCGGCTCTCTCAAAAGAGACCCCTTACCACTGGCTCTGATGAAGAG	364
02886	265	GGGTCACTGGGGCCCATGGATGTCGTTGCTTAACCTGAGTGCCTTTTGGTGGAGAGGCGCGGCTCTCTCAAAAGAGACCCCTTACCACTGGCTCTGATGAAGAG	364
03886	265	GGGTCACTGGGGCCCATGGATGTCGTTGCTTAACCTGAGTGCCTTTTGGTGGAGAGGCGCGGCTCTCTCAAAAGAGACCCCTTACCACTGGCTCTGATGAAGAG	350
04886	251	GGGTCACTGGGGCCCATGGATGTCGTTGCTTAACCTGAGTGCCTTTTGGTGGAGAGGCGCGGCTCTCTCAAAAGAGACCCCTTACCACTGGCTCTGATGAAGAG	367
05882	268	GGGTCACTGGGGCCCATGGATGTCGTTGCTTAACCTGAGTGCCTTTTGGTGGAGAGGCGCGGCTCTCTCAAAAGAGACCCCTTACCACTGGCTCTGATGAAGAG	366
06886	267	GGGTCACTGGGGCCCATGGATGTCGTTGCTTAACCTGAGTGCCTTTTGGTGGAGAGGCGCGGCTCTCTCAAAAGAGACCCCTTACCACTGGCTCTGATGAAGAG	365
08886	266	GGGTCACTGGGGCCCATGGATGTCGTTGCTTAACCTGAGTGCCTTTTGGTGGAGAGGCGCGGCTCTCTCAAAAGAGACCCCTTACCACTGGCTCTGATGAAGAG	365
12886	266	GGGTCACTGGGGCCCATGGATGTCGTTGCTTAACCTGAGTGCCTTTTGGTGGAGAGGCGCGGCTCTCTCAAAAGAGACCCCTTACCACTGGCTCTGATGAAGAG	365
13886	266	GGGTCACTGGGGCCCATGGATGTCGTTGCTTAACCTGAGTGCCTTTTGGTGGAGAGGCGCGGCTCTCTCAAAAGAGACCCCTTACCACTGGCTCTGATGAAGAG	370
14886	271	GGGTCACTGGGGCCCATGGATGTCGTTGCTTAACCTGAGTGCCTTTTGGTGGAGAGGCGCGGCTCTCTCAAAAGAGACCCCTTACCACTGGCTCTGATGAAGAG	320
19886	221	GGGTCACTGGGGCCCATGGATGTCGTTGCTTAACCTGAGTGCCTTTTGGTGGAGAGGCGCGGCTCTCTCAAAAGAGACCCCTTACCACTGGCTCTGATGAAGAG	367
22886	268	GGGTCACTGGGGCCCATGGATGTCGTTGCTTAACCTGAGTGCCTTTTGGTGGAGAGGCGCGGCTCTCTCAAAAGAGACCCCTTACCACTGGCTCTGATGAAGAG	367
008TD	401	GAGTACACAGAAACATATAATTCAGGAAGCAGCTTTCCCAATGCTCGACATCATCCAGGCCAATCCCGCTCTCTGGTTCTCTCCCTCTCTCTCTGGACT	500
01886	365	GAGTACACAGAAACATATAATTCAGGAAGCAGCTTTCCCAATGCTCGACATCATCCAGGCCAATCCCGCTCTCTGGTTCTCTCCCTCTCTCTCTGGACT	464
02886	365	GAGTACACAGAAACATATAATTCAGGAAGCAGCTTTCCCAATGCTCGACATCATCCAGGCCAATCCCGCTCTCTGGTTCTCTCCCTCTCTCTCTGGACT	464
03886	365	GAGTACACAGAAACATATAATTCAGGAAGCAGCTTTCCCAATGCTCGACATCATCCAGGCCAATCCCGCTCTCTGGTTCTCTCCCTCTCTCTCTGGACT	464
04886	351	GAGTACACAGAAACATATAATTCAGGAAGCAGCTTTCCCAATGCTCGACATCATCCAGGCCAATCCCGCTCTCTGGTTCTCTCCCTCTCTCTCTGGACT	450
05882	368	GAGTACACAGAAACATATAATTCAGGAAGCAGCTTTCCCAATGCTCGACATCATCCAGGCCAATCCCGCTCTCTGGTTCTCTCCCTCTCTCTCTGGACT	467
06886	367	GAGTACACAGAAACATATAATTCAGGAAGCAGCTTTCCCAATGCTCGACATCATCCAGGCCAATCCCGCTCTCTGGTTCTCTCCCTCTCTCTCTGGACT	466
08886	366	GAGTACACAGAAACATATAATTCAGGAAGCAGCTTTCCCAATGCTCGACATCATCCAGGCCAATCCCGCTCTCTGGTTCTCTCCCTCTCTCTCTGGACT	465
12886	366	GAGTACACAGAAACATATAATTCAGGAAGCAGCTTTCCCAATGCTCGACATCATCCAGGCCAATCCCGCTCTCTGGTTCTCTCCCTCTCTCTCTGGACT	465
13886	366	GAGTACACAGAAACATATAATTCAGGAAGCAGCTTTCCCAATGCTCGACATCATCCAGGCCAATCCCGCTCTCTGGTTCTCTCCCTCTCTCTCTGGACT	465
14886	371	GAGTACACAGAAACATATAATTCAGGAAGCAGCTTTCCCAATGCTCGACATCATCCAGGCCAATCCCGCTCTCTGGTTCTCTCCCTCTCTCTCTGGACT	470
19886	321	GAGTACACAGAAACATATAATTCAGGAAGCAGCTTTCCCAATGCTCGACATCATCCAGGCCAATCCCGCTCTCTGGTTCTCTCCCTCTCTCTCTGGACT	420
22886	368	GAGTACACAGAAACATATAATTCAGGAAGCAGCTTTCCCAATGCTCGACATCATCCAGGCCAATCCCGCTCTCTGGTTCTCTCCCTCTCTCTCTGGACT	467
008TD	501	CTGTGCACACGCTCTCTCTCTCTGAGGCTGAATTT	533
01886	465	CTGTGCACACGCTCTCTCTCTCTGAGGCTGAATTT	497
02886	465	CTGTGCACACGCTCTCTCTCTCTGAGGCTGAATTT	497
03886	465	CTGTGCACACGCTCTCTCTCTCTGAGGCTGAATTT	497
04886	451	CTGTGCACACGCTCTCTCTCTCTGAGGCTGAATTT	483
05882	468	CTGTGCACACGCTCTCTCTCTCTGAGGCTGAATTT	500
06886	467	CTGTGCACACGCTCTCTCTCTCTGAGGCTGAATTT	499
08886	466	CTGTGCACACGCTCTCTCTCTCTGAGGCTGAATTT	498
12886	466	CTGTGCACACGCTCTCTCTCTCTGAGGCTGAATTT	498
13886	466	CTGTGCACACGCTCTCTCTCTCTGAGGCTGAATTT	498
14886	471	CTGTGCACACGCTCTCTCTCTCTGAGGCTGAATTT	503
19886	421	CTGTGCACACGCTCTCTCTCTCTGAGGCTGAATTT	453
22886	468	CTGTGCACACGCTCTCTCTCTCTGAGGCTGAATTT	500

FIGURE 11B (continued)

# SEQUENCE LISTING

<110> Abreu, Maria T.  
Taylor, Kent D.  
Rotter, Jerome I.  
Yang, Huiying  
Sugimura, Kazuhito  
Targan, Stephan R.

<120> Mutations in NOD2 are Associated with  
Fibrostenosing Disease in Patients with Crohn's Disease

<130> P-CE 5480

<150> US 60/407,391

<151> 2002-08-30

<160> 67

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 494

<212> DNA

<213> Homo sapiens

<400> 1

```
accttcagat cacagcagcc ttcttggcag ggctgttgtc ccgggagcac tggggcctgc 60
tggctgagtg ccagacatct gagaaggccc tgctccggcg ccaggcctgt gcccgctggg 120
gcttggcccg cagcctccgc aagcacttcc actccatccc gccagctgca ccgggtgagg 180
ccaagagcgt gcatgccatg cccgggttca tctggctcat ccggagcctg tacgagatgc 240
aggaggagcg gctggctcgg aaggctgcac gtggcctgaa tgttgggcac ctcaaagtga 300
cattttgcag tgtgggcccc actgagtgtg ctgccctggc ctttgtgctg cagcacctcc 360
ggcggcccggt ggccctgcag ctggactaca actctgtggg tgacattggc ctggagcagc 420
tgctgccttg ccttgggtgc tgcaaggctc tgtagtgagt gttactgggc attgctgttc 480
aggtatgggg gagc 494
```

<210> 2

<211> 494

<212> DNA

<213> Homo sapiens

<400> 2

```
gctcccccat acctgaacag caatgccag taacactcac tacagagcct tgcagacacc 60
aaggcaaggc agcagctgct ccaggccaat gtcacccaca gagttgtagt ccagctgcag 120
ggccacgggc cgccggaggt gctgcagcac aaaggccagg gcagcacact cagtggggcc 180
cacactgcaa aatgtcaact tgagggtgcc aacattcagg ccacgtgcag ccttccgagc 240
cagccgctcc tctgcatct cgtacagget ccggatgagc cagatgaacc cgggcatggc 300
atgcacgctc ttggcctcac ccggtgcagc tggcgggatg gagggaagt gcttgcggag 360
gctgcggggc agacaccagc gggcacaggc ctggcgccgg agcagggcct tctcagatgt 420
```

ctggcactca gccagcaggc cccagtgtc cggggacaac agccctgcc ggaaggctgc 480  
tgtgatctga aggt 494

<210> 3  
<211> 540  
<212> DNA  
<213> Homo sapiens

<400> 3  
atcaaaaccc tgagaggaca agggacattt ccaagtcacc cagaaagact cgagtgtcct 60  
ctcttgaaat ccaatgggtct tttttcctta ctccattgcc taacattgtg gggtagaaat 120  
aaagttcaaa gaccttcaga actggcccca gtcctccct ctccacctga tctccccaag 180  
aaaactgcag gatagactct gaagettacc tgagccacct caagctctgg tgatcaccca 240  
aggcttcagc cagggcctgg gcccctcgt caccactct gttgcccag aatctgaaaa 300  
ggccaaaaga gtcaacagac agtgtcagtg agtacctgat atgtgttcta gacatgaact 360  
aacagtcctc ctccctctgc agtcccagcc agaggggag gaccactcaa tcccagagt 420  
gcctcactgg ggctcctggt cccagcaaaag tggacctgcc tccatctttt ggggtgggatg 480  
gccaaactta acccaagagt tttcagtggt tttacattac agacttagag aatagtagag 540

<210> 4  
<211> 540  
<212> DNA  
<213> Homo sapiens

<400> 4  
ctctactatt ctctaagtct gtaatgtaaa gccactgaaa actcttgggt taagtttggc 60  
catcccaccc aaaagatgga ggcaggtcca ctttgctggg accaggagcc ccagtgaggc 120  
caactctggga ttgagtggtc ctgcccctct ggctgggact gcagagggag gaggactgtt 180  
agttcatgtc tagaacacat atcaggtact cactgacact gtctgttgac tcttttggcc 240  
ttttcagatt ctggggcaac agagtgggtg acgagggggc ccaggccctg gctgaagcct 300  
tgggtgatca ccagagcttg aggtggctca ggtaagcttc agagtctatc ctgcagtttt 360  
cttggggaga tcaggtgaag agggaggagc tggggccagt tctgaaggtc tttgaacttt 420  
atctctaccc cacaatgtta ggcaatggag taaggaaaaa agaccattgg atttcaagag 480  
aggacactcg agtctttctg ggtgacttgg aaatgtccct tgtcctctca gggttttgat 540

<210> 5  
<211> 541  
<212> DNA  
<213> Homo sapiens

<400> 5  
tttaaaaatg aaatcattgc tccctactta aagaggtaaa gacttctttc ttagacagag 60  
aatcagatcc ttcacatgca gaatcattct cactgaatgt cagaatcaga agggatcctc 120  
aaaattctgc cattcctctc tcccgtcacc ccattttaca gatagaaaaa ctgagggttcg 180  
gagagctaaa acaggcctgc ccaggggctt taccagactt ccaggatggg gtcattcctt 240  
tcaagggggc tgcaggaggg cttctgcccc taggtagggt atgcagttat tggacaacct 300  
ggaaaagaag atacaatggt gagcttcaag gattcttggg tttcctcttg aaactgtcca 360  
gttaaaagaga ctgcaggagt tagccagtct actgaagccc acctgtccct tagacacatc 420  
ctgctcatgt ctgagattcc caatgagtc atcaacaaaag gctcagtagc atcagtgaaa 480  
tgtaaccgtc tctcttccat tcactagatg agtttatcaa attaaagtag cactccctta 540  
g 541

- 3 -

<210> 6  
 <211> 541  
 <212> DNA  
 <213> Homo sapiens

<400> 6  
 ctaaggaggagt ggctacttaa tttgataaac tcatctagtg aatggaagag agacgggttac 60  
 atttcactga tggtagctag cctttgttga tgagctcatt gggaaatctca gacatgagca 120  
 ggatgtgtct aaggacaggg tgggcttcag tagactggct aactcctgca gtctctttaa 180  
 ctggacagtt tcaagaggaa aaccaagaat ccttgaagct caccattgta tcttcttttc 240  
 cagggtgtcc aataactgca tcacctacct aggggcagaa gccctcctgc agggcccttg 300  
 aaaggaatga caccatcctg gaagtctggt aaggccctg ggcaggcctg ttttagctct 360  
 ccgaacctca gtttttctat ctgtaaaatg gggtagcggg agagaggaat ggcagaattt 420  
 tgaggatccc ttctgattct gacattcagt gagaatgatt ctgcatgtga aggatctgat 480  
 tctctgtcta agaaagaagt ctttacctct ttaagtaggg agcaatgatt tcatttttaa 540  
 a 541

<210> 7  
 <211> 540  
 <212> DNA  
 <213> Homo sapiens

<400> 7  
 aacagcagtg ctcaaagagt agagtccgca cagagagtggt tttggccatg cactgcagct 60  
 gccggcagct gaatgggaag acaaagagaa attcctggaa gtcttgccct gcagcccaca 120  
 gcaagtgcag ccgctgcagg agcgtgctct tgccactgcc cgctcacc accaccagca 180  
 cagtgtccgc atcgtcattg aggtggccag gggtagtgaa gagctcctcc agggccaggg 240  
 tggctgggct cttctgcggg ggtccagcca tgcccacatc tgcccagacc tccaggacat 300  
 tctctgtgta tatgtcctcc aggcagagcg tctctgctcc atcataggta ctgaggaagc 360  
 gagactgagc agacaccgtg gtcctcagct tggccatata cttcttgcat gtggcagctg 420  
 gaaggcagaa gaagaggcag atgaagggtg caccatgggt aagacgggac ctaaccagac 480  
 aatgggctgc tgcgggggac gctgacataa ctgaagggtat aggagagcca gcggggcgccc 540

<210> 8  
 <211> 540  
 <212> DNA  
 <213> Homo sapiens

<400> 8  
 gggcgcccg c tggctctcct atcccttcag ttatgtcagc gtcccccga gcagcccatt 60  
 gtctgggttag gtcccgctct caccatgggt ccaccttcat ctgcctcttc ttctgccttc 120  
 cagctgccac atgcaagaag tatatggcca agctgaggac cagggtgtct gctcagctctc 180  
 gcttcctcag tacctatgat ggagcagaga cgctctgcct ggaggacata tacacagaga 240  
 atgtcctgga ggtctgggca gatgtgggca tggctggacc cccgcagaag agcccagcca 300  
 ccctgggcct ggaggagctc ttcagcacc ctaggcacct caatgacgat gcggacactg 360  
 tgctgggtggg gggtagggcg ggcagtggca agagcacgct cctgcagcgg ctgcacttgc 420  
 tgtgggctgc agggcaagac ttccaggaat ttctctttgt cttccattc agctgccggc 480  
 agctgcagtg catggccaaa ccactctctg tgcggactct actctttgag cactgctggt 540

<210> 9

<211> 520  
<212> DNA  
<213> Homo sapeins

<400> 9  
gcactgggca cccactacca atggattgga attggctcctt aagataaaat gtacctgac 60  
cagcccaata tcttcaattt acagatactg tatcaaaacc ctgagaggac aaggacatt 120  
tccaagtcac ccagaaagac tcgagtgtcc tctcttgaaa tccaatgggc ttttttcctt 180  
actccattgc ctaacattgt ggggtagaaa taaagttcaa agaccttcag aactggcccc 240  
agctcctccc tcttcacctg atctcccca gaaaactgca ggatagactc tgaagcttac 300  
ctgagccacc tcaagctctg gtgatcacc aaggcttcag ccagggcctg ggccccctcg 360  
tcacccactc tgttgcccca gaatctgaaa aggccaaaag agtcaacaga cagtgtcagt 420  
gagtacctga tatgtgttct agacatgaac taacagtcct cctccctctg cagtcccagc 480  
cagaggggca ggaccactca atcccagagt ggctcactg 520

<210> 10  
<211> 520  
<212> DNA  
<213> Homo sapiens

<400> 10  
cagtggaggcc actctgggat tgagtgggtcc tgcccctctg gctgggactg cagagggagg 60  
aggactgtta gttcatgtct agaacacata tcaggctactc actgacactg tctgttgact 120  
cttttggcct tttcagattc tggggcaaca gagtgggtga cgagggggcc caggccctgg 180  
ctgaagcctt ggggtgatcac cagagcttga ggtggctcag gtaagcttca gagtctatcc 240  
tgcagtttct ttggggagat cagggtgaaga gggaggagct ggggccagtt ctgaaggctc 300  
ttgaacttta tttctacccc acaatgttag gcaatggagt aaggaaaaaa gaccattgga 360  
tttcaagaga ggacactcga gtctttctgg gtgacttgga aatgtccctt gtcctctcag 420  
ggttttgata cagtatctgt aaattgaaga tattgggctg gatcaggtac attttatctt 480  
aaggaccaat tccaatccat tggtagtggg tgcccagtgc 520

<210> 11  
<211> 535  
<212> DNA  
<213> Homo sapiens

<400> 11  
tcactaacca gctcaggaag ctcaccagct tgggaagtta atcattatgt ctgacttcag 60  
tttctcctgc ttcagtttaa attgggaaag agagagaaaa aatattcact cattatctgt 120  
ttcctaaaaat tgtccttaac atccttcctc ttactccttt attacctggg cgggcttccc 180  
ctcttcaggc gaaatctgtc agtctatctg cattgccttt tgatctctac ttcagttact 240  
acaacttcaa agacaccatt gtcctcccca aggtgaggcc catgtagaga aaggatcact 300  
tccttgctga aagagagggt caaggggcga cccacgtggg cctccctga aaccaggcc 360  
caggcctgag cctggacacc tccttccttc ctgagaccac agccagcccg gtttctctgg 420  
ggccaagagc aaatgctttg ctttaagtgt gaaatctcag cccactgacc ccttgccagc 480  
aggagaggag ggggaagcca gggaaagctca acttcccaag tgtcctgagt ctctg 535

<210> 12  
<211> 496  
<212> DNA  
<213> Homo sapiens

<400> 12

- 5 -

```

aatcattatg tctagcttca gtttctcctg cttcagttta aattgggaaa gagagagaaa 60
aaatattcac tcattatctg ttctctaaaa ttgtccttaa catccttcct cttactcctt 120
tattacctgg tcgggcttcc cctcttcagg cgaaatctgt cagtctatct gcattgcctt 180
ttgatctcta cttcagttac tacmacttca aagacaccat tgtcctcccc aaggtagggs 240
ccatgtagag aaaggatcac ttccttgctg aaagagaggg tcaaggggtg acccacgtgg 300
gccctccctg aaaccagggc ccaggcctga gcctggacac ctcttcctt cctgagacca 360
cagccagccc ggtttctctg gggccaagag caaatgcttt gcttaagtgc tgaaatctca 420
gccactgac cccttgcmga caggagagga ggggaagccc agggaagctc aacttcccaa 480
gtgtcctgag tctctg                                     496

```

&lt;210&gt; 13

&lt;211&gt; 488

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 434

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 13

```

tgtctagctt cagtttctcc tgccttcagtt taaattggga aagagagaga aaaaatattc 60
actcattatc tgtttcctaa aattgtcctt aacatccttc ctcttactcc tttattacct 120
ggtcgggctt cccctcttca ggcgaaatct gtcagtctat ctgcattgcc ttttgatctc 180
tacttcagtt actacaactt caaagacacc attgtcctcc ccaaggtagg gcccatgtag 240
agaaaggatc acttccttgc tgaaagagag ggtcaagggg tgaccacagt gggccctccc 300
tgaaaccagg gccaggcct gagcctggac acctccttcc ttcctgagac cacagccagc 360
ccggtttctc tggggccaag agcaaatgct ttgcttaagt gctgaaatct cagccactg 420
amcccttgca gacnggagag gaggggaagc ccagggaagc tcaacttccc aagtgtcctg 480
agtctctg                                     488

```

&lt;210&gt; 14

&lt;211&gt; 491

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 437

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 14

```

ttatgtctag cttcagtttc tctgcttca gtttaaattg ggaaagagag agaaaaaata 60
ttcactcatt atctggttcc taaaattgtc cttaacatcc ttctctttac tcctttatta 120
cctgggtcggg cttccctctc tcaggcgaaa tctgtcagtc tatctgcatt gccttttgat 180
ctctacttca gttactacaa cttcaaagac accattgtcc tccccaaggt gaggcccatg 240
tagagaaagg atcacttcct tgctgaaaga gaggggtcaag ggyygacca cgtgggccc 300
ccctgaaacc caggcccagg cctgagcctg gacacctcct tccttcctga gaccacagcc 360
agcccgggtt ctctggggcc aagagcaaat gctttgctta agtgctgaaa tctcagccca 420
ctgacccctt gcagacngga gaggagggga agcccaggga agctcaactt cccaagtgtc 480
ctgagtctct g                                     491

```

&lt;210&gt; 15

- 6 -

<211> 491  
 <212> DNA  
 <213> Homo sapiens

<400> 15  
 ttatgtctag cttcagtttc tctgtcttca gtttaaattg ggaaagagag agaaaaaata 60  
 ttcaactcatt atctgtttcc taaaattgtc cttaacatcc ttctctttac tcctttatta 120  
 cctggtcggg cttccctctc tcaggcgaaa tctgtcagtc tatctgcatt gccttttgat 180  
 ctctacttca gttactacaa cttcaaagac accattgtcc tccccaagggt gagggccatg 240  
 tagagaaagg atcacttcct tgctgaaaga gaggggtcaag ggygaccca cgtgggcccct 300  
 ccctgaaacc caggcccagg cctgagcctg gacacctcct tccttctga gaccacagcc 360  
 agcccggttt ctctggggcc aagagcaaat gctttgctta agtgctgaaa tctcagccca 420  
 ctgacccctt gcagacagga gaggagggga agcccagggga agctcaactt cccaagtgtc 480  
 ctgagtctct g 491

<210> 16  
 <211> 491  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 437  
 <223> n = A,T,C or G

<400> 16  
 ttatgtctag cttcagtttc tctgtcttca gtttaaattg ggaaagagag agaaaaaata 60  
 ttcaactcatt atctgtttcc taaaattgtc cttaacatcc ttctctttac tcctttatta 120  
 cctggtcggg cttccctctc tcaggcgaaa tctgtcagtc tatctgcatt gccttttgat 180  
 ctctacttca gttactacaa cttcaaagac accattgtcc tccccaagggt gagggccatg 240  
 tagagaaagg atcacttcct tgctgaaaga gaggggtcaag ggygaccca cgtgggcccct 300  
 ccctgaaacc caggcccagg cctgagcctg gacacctcct tccttctga gaccacagcc 360  
 agcccggttt ctctggggcc aagagcaaat gctttgctta agtgctgaaa tctcagccca 420  
 ctgacccctt gcagacngga gaggagggga agcccagggga agctcaactt cccaagtgtc 480  
 ctgagtctct g 491

<210> 17  
 <211> 491  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 159  
 <223> n = A,T,C or G

<400> 17  
 ttatgtctag cttcagtttc tctgtcttca gtttaaattg ggaaagagag agaaaaaata 60  
 ttcaactcatt atctgtttcc taaaattgtc cttaacatcc ttctctttac tcctttatta 120  
 cctggtcggg cttccctctc tcaggcgaaa tctgtcagnc tatctgcatt gccttttgat 180  
 ctctacttca gttactacaa cttcaaagac accattgtcc tccccaagggt gagggccatg 240  
 tagagaaagg atcacttcct tgctgaaaga gaggggtcaag ggygaccca cgtgggcccct 300  
 ccctgaaacc caggcccagg cctgagcctg gacacctcct tccttctga gaccacagcc 360

- 7 -

agcccggttt ctctggggcc aagagcaaatt gctttgctta agtgctgaaa tctcagccca 420  
 ctgacccctt gcagacagga gaggagggga agcccaggga agtcaactt cccaagtgtc 480  
 ctgagtctct g 491

<210> 18  
 <211> 487  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 433  
 <223> n = A,T,C or G

<400> 18  
 gtctagcttc agtttctcct gcttcagttt aaattgggaa agagagagaa aaaatattca 60  
 ctyattatct gtttctctaa attgtcctta acatccttcc tcttactcct ttattacctg 120  
 gtcgggcttc cctcttcag gcgaaatctg tcagtctatc tgcattgcct ttgatctct 180  
 acttcagtta ctacaacttc aaagacacca ttgtcctccc caagggtgagg cccatgtaga 240  
 gaaaggatca cttccttgct gaaagagagg gtcaaggggy gaccacgtg ggccctccct 300  
 gaaaccagg cccaggcctg agcctggaca cctccttcc tctgagacc acagccagcc 360  
 cggtttctct ggggccaaga gcaaatgctt tgettaagt ctgaaatctc agccactga 420  
 ccccttgtag acnggagagg aggggaagcc cagggaagct caacttccca agtgtcctga 480  
 gtctctg 487

<210> 19  
 <211> 486  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 432  
 <223> n = A,T,C or G

<400> 19  
 tctagcttca gtttctcctg cttcagttta aattgggaaa gagagagaaa aaatattcac 60  
 tcattatctg tttcctaaaa ttgtccttaa catccttccct cttactcctt tattacctgg 120  
 tcgggcttcc cctcttcagg cgaaatctgt cagtctatct gcattgcctt ttgatctcta 180  
 cttcagttac tacaacttca aagacaccat tgtcctcccc aagggtgaggc ccatgtaga 240  
 aaaggatcac ttccttgctg aaagagaggg tcaaggggcy acccacgtgg gccctccctg 300  
 aaaccaggc ccaggcctga gcttgacac ctccttccct cctgagacca cagccagccc 360  
 gggttctctg gggccaagag caaatgctt gcttaagtgc tgaaatctca gccactgac 420  
 cccttgtaga cnggagagga ggggaagccc aggggaagctc aacttcccaa gtgtcctgag 480  
 tctctg 486

<210> 20  
 <211> 484  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature



- 8 -

&lt;222&gt; 430

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 20

```

tagcttcagt ttctcctgct tcagtttaaa ttgggaaaga gagagaaaaa atattcactc 60
attatctggt tcctaaaatt gtccttaaca tccttcctct tactccttta ttacctggtc 120
gggcttcccc tcttcaggcg aaatctgtca gtctatctgc attgcctttt gatctctact 180
tcagttacta caacttcaaa gacaccattg tcctcccaa ggtgaggccc atgtagagaa 240
aggatcactt ccttgctgaa agagagggtc aaggggagac ccacgtgggc cctccctgaa 300
acccaggccc aggcctgagc ctggacacct ccttccttcc tgagaccaca gccagcccg 360
tttctctggg gccaaagaca aatgctttgc ttaagtgtg aaatctcagc ccactgacct 420
cttgacagac ggagaggagg ggaagcccag ggaagctcaa cttcccaagt gtcctgagtc 480
tctg 484

```

&lt;210&gt; 21

&lt;211&gt; 485

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 431

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 21

```

ctagcttcag tttctcctgc ttcagtttaa attgggaaag agagagaaaa aatattcact 60
yattatctgt ttctaaaaat tgtccttaac atccttcctc ttactccttt attacctggt 120
cgggcttccc ctcttcaggc gaaatctgtc agtctatctg cattgccttt tgatctctac 180
ttcagttact acaacttcaa agacaccatt gtcctcccaa aggtgaggcc catgtagaga 240
aaggatcact tccttgctga aagagagggt caaggggaga cccacgtggg ccctccctga 300
aaccaggccc caggcctgag cctggacacc tccttccttc ctgagaccac agccagcccg 360
gtttctctgg ggccaagagc aaatgctttg cttaagtgtg gaaatctcag cccactgacc 420
ccttgacagc nggagaggag ggaagcccaa ggaagctca acttcccaag tgtcctgagt 480
ctctg 485

```

&lt;210&gt; 22

&lt;211&gt; 488

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 434

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 22

```

tgtctagctt cagtttctcc tgettcagtt taaattggga aagagagaga aaaaatattc 60
acttattatc tgtttcctaa aattgtcctt aacatccttc ctcttactcc tttattacct 120
ggtcgggctt cccctcttca ggcgaaatct gtcagtctat ctgcattgcc ttttgatctc 180
tacttcagtt actacaactt caaagacacc attgtcctcc ccaaggtag gcccattgag 240
agaaaggatc acttccttgc tgaaagagag ggtcaagggg cgacccacgt gggccctccc 300
tgaaacccag gccagggcct gagcctggac acctccttcc ttcttgagac cacagccagc 360
ccggtttctc tggggccaag agcaaagtct ttgcttaagt gctgaaatct cagcccactg 420

```

- 9 -

acccttgcac gacnngagag gaggggaagc ccaggggaagc tcaacttccc aagtgtcctg 480  
agtctctg 488

<210> 23  
<211> 488  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<222> 434  
<223> n = A,T,C or G

<400> 23  
tgtctagctt cagtttctcc tgcttcagtt taaattggga aagagagaga aaaaatattc 60  
acttattatc tgtttctctaa aattgtcctt aacatccttc ctcttactcc tttattacct 120  
ggctgggctt cccctcttca ggcgaaatct gtcagtctat ctgcattgcc ttttgatctc 180  
tacttcagtt actacaactt caaagacacc attgtcctcc ccaagggtgag gcccatgtag 240  
agaaaggatc acttccttgc tgaaagagag ggtcaagggg cgacccacgt gggccctccc 300  
tgaaacccag gccaggcct gagcctggac acctccttcc ttcttgagac cacagccagc 360  
ccggtttctc tggggccaag agcaaatgct ttgcttaagt gctgaaatct cagcccactg 420  
acccttgcac gacnngagag gaggggaagc ccaggggaagc tcaacttccc aagtgtcctg 480  
agtctctg 488

<210> 24  
<211> 497  
<212> DNA  
<213> Homo sapiens

<400> 24  
tcactaggct tctggttgat gcctgtgaac tgaactctga caacagactt ctgaaataga 60  
cccacaagag gcagttccat ttcatttctg ccagaatgct ttaggatgta cagttatgga 120  
ttgaaagttt acaggaaaaa aaattaggcc gttccttcaa agcaaatgtc ttcttgatt 180  
attcaaaatg atgtatgttg aagcctttgt aaattgtcag atgctgtgca aatgttatta 240  
ttttaaacat tatgatgtgt gaaaactggg taatatttat aggtcacttt gttttactgt 300  
cttaagttta tactcttata gacaacatgg ccgtgaactt tatgctgtaa ataacagag 360  
gggaataaac tggtgagtc aaacagccat cttccttggtg accaaacatt taaaaatatt 420  
ctggctgggc acagtggctc acgcctgtaa tcccagcact ttggggaggcc gaggtgggca 480  
gatcacctga ggttggg 497

<210> 25  
<211> 460  
<212> DNA  
<213> Homo sapiens

<400> 25  
tgacaacaga cttctgaaat agaccacaaa gaggcagttc catttcattt gtgccagaat 60  
gctttaggat gtacagttat ggattgaaag ttacaggaa aaaaaattag gccgttccct 120  
caaagcaaat gtcttccctg attattcaaa atgatgtatg ttgaagcctt tgtaaatgt 180  
cagatgctgt gcaaatgtta ttattttaaa cattatgatg tgtgaaaact ggttaatat 240  
tataggtcac tttgttttac tgtcttaagt ttatactctt atagacaaca tggccgtgaa 300  
ctttatgctg taaataatca gaggggaata aactgttgag tcaaaacagc catcttccct 360  
gtgaccaaac atttaaaat attctggctg ggcacagtgg ctacgcctg taatcccagc 420

- 10 -

actttgggag gccgaggtgg gcagatcacc tgaggttggg

460

&lt;210&gt; 26

&lt;211&gt; 462

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 26

tctgacaaca gacttctgaa atagacccac aagaggcagt tccatttcat ttgtgccaga 60  
 atgcttttagg atgtacagtt atggattgaa agtttacagg aaaaaaaatt aggccgttcc 120  
 ttcaaagcaa atgtcttcct ggattattca aaatgatgta tgttgaagcc ttgttaaatt 180  
 gtcagatgct gtgcaaatgt tattatttta aacattatga tgtgtgaaaa ctggttaata 240  
 tttataggtc actttgtttt actgtcttaa gtttatactc ttatagacaa catggccgtg 300  
 aactttatgc tgtaataaat cagaggggaa taaactgttg agtcaaaaca gccatcttcc 360  
 ttgtgaccaa acatttataa atattctggc tgggcacagt ggctcacgcc tgtaatccca 420  
 gcactttggg aggccgaggt gggcagatca cctgaggttg gg 462

&lt;210&gt; 27

&lt;211&gt; 459

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 27

gacaacagac ttctgaaata gaccacaag aggcagttcc atttcatttg tgccagaatg 60  
 ctttaggatg tacagttatg gattgaaagt ttacaggaaa aaaaattagg ccgttccttc 120  
 aaagcaaatg tcttcctgga ttattcaaaa tgatgtatgt tgaagccttt gtaaattgtc 180  
 agatgctgtg caaatgttat tatttttaaac attatgatgt gtgaaaactg gttaatatgt 240  
 ataggtcact ttgttttact gtcttaagtt tatactctta tagacaacat ggccgtgaac 300  
 tttatgctgt aaataatcag aggggaataa actgttgagt caaaacagcc atcttccttg 360  
 tgaccaaaaa tttaaaaata ttctggctgg gcacagtggc tcacgcctgt aatcccagca 420  
 ctttgggagg ccgaggtggg cagatcacct gaggttggg 459

&lt;210&gt; 28

&lt;211&gt; 467

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 28

tgaactctga caacagactt ctgaaataga cccacaagag gcagttccat ttcatttgtg 60  
 ccagaatgct ttaggatgta cagttatgga ttgaaagttt acaggaaaaa aaattaggcc 120  
 gttccttcaa agcaaatgte ttcttgatt attcaaaatg atgtatgttg aagcctttgt 180  
 aaattgtcag atgctgtgca aatgttatta ttttaaacat tatgatgtgt gaaaactggg 240  
 taatatattat aggtcacttt gttttactgt cttaagttta tactcttata gacaacatgg 300  
 ccgtgaactt tatgctgtaa ataatcagag gggaataaac tgttgagtca aaacagccat 360  
 cttccttgtg accaaacatt taaaaatatt ctggctgggc acagtggctc acgcctgtaa 420  
 tcccagcact ttgggagggc gaggtgggca gatcacctga gggtggg 467

&lt;210&gt; 29

&lt;211&gt; 467

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 29

- 11 -

```

tgaactctga caacagactt ctgaaataga cccacaagag gcagttccat ttcattttgtg 60
ccagaatgct ttaggatgta cagttatgga ttgaaagttt acaggaaaaa aaattaggcc 120
gttccttcaa agcaaatgtc ttcttggtt attcaaaatg atgtatgttg aagcctttgt 180
aaattgtcag atgctgtgca aatgttatta ttttaaacat tatgatgtgt gaaaactggt 240
taatatttat agrtcacttt gttttactgt cttaagttaa tactcttata gacaacatgg 300
ccgtgaactt tatgctgtaa ataatacagag gggaataaaac tgttgagtca aaacagccat 360
cttccttgtg accaaacatt taaaaatatt ctggctgggc acagtggctc acgcctgtaa 420
tcccagcact ttgggaggcc gaggtgggca gatcacctga ggttggg 467

```

&lt;210&gt; 30

&lt;211&gt; 466

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 30

```

gaactatgac aacagacttc tgaaatagac ccacaagagg cagttccatt tcattttgtgc 60
cagaatgctt taggatgtac agttatggat tgaaagttaa caggaaaaaa aattaggccg 120
ttccttcaaa gcaaatgtct tcctggatta ttcaaaatga tgtatgttga agcctttgta 180
aattgtcaga tgctgtgcaa atgttattat ttttaaacatt atgatgtgtg aaaactggtt 240
aatatttata grtcactttg ttttactgtc ttaagttaa actcttatag acaacatggc 300
cgtgaacttt atgctgtaaa taatcagagg ggaataaaact gttgagtcaa aacagccatc 360
ttccttgtga ccaaacattt aaaaatattc tggctgggca cagtggctca cgctgtaat 420
cccagcactt tgggaggccg aggtgggcag atcacctgag gttggg 466

```

&lt;210&gt; 31

&lt;211&gt; 466

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 31

```

gaactctgac aacagacttc tgaaatagac ccacaagagg cagttccatt tcattttgtgc 60
cagaatgctt taggatgtac agttatggat tgaaagttaa caggaaaaaa aattaggccg 120
ttccttcaaa gcaaatgtct tcctggatta ttcaaaatga tgtatgttga agcctttgta 180
aattgtcaga tgctgtgcaa atgttattat ttttaaacatt atgatgtgtg aaaactggtt 240
aatatttata grtcactttg ttttactgtc ttaagttaa actcttatag acaacatggc 300
cgtgaacttt atgctgtaaa taatcagagg ggaataaaact gttgagtcaa aacagccatc 360
ttccttgtga ccaaacattt aaaaatattc tggctgggca cagtggctca cgctgtaat 420
cccagcactt tgggaggccg aggtgggcag atcacctgag gttggg 466

```

&lt;210&gt; 32

&lt;211&gt; 460

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 32

```

tgacaacaga cttctgaaat agaccacaaa gaggcagttc catttcattt gtgccagaat 60
gcttttaggat gtacagttat ggattgaaag ttacaggaa aaaaaattag gccgttcctt 120
caaagcaaat gtcttctggt attattcaaa atgatgtatg ttgaagcctt tgtaaattgt 180
cagatgctgt gcaaatgtta ttattttaaa cattatgatg tgtgaaaact ggtaaatatt 240
tatagrtcac tttgttttac tgtcttaagt ttatactctt atagacaaca tggccgtgaa 300
ctttatgctg taaataatca gaggggaata aactgttgag tcaaaacagc catcttcctt 360
gtgaccaaac atttaaaaat attctggctg ggcacagtgg ctcacgcctg taatccagc 420
actttgggag gccgaggtgg gcagatcacc tgaggttggg 460

```

- 12 -

<210> 33  
 <211> 467  
 <212> DNA  
 <213> Homo sapiens

<400> 33  
 tgaactctga caacagactt ctgaaataga cccacaagag gcagttccat ttcatttgtg 60  
 ccagaatgct ttaggatgta cagttatgga ttgaaagttt acaggaaaaa aaattaggcc 120  
 gttccttcaa agcaaagtgc ttcctggatt attcaaaatg atgtatgttg aagcctttgt 180  
 aaattgtcag atgctgtgca aatggttatta ttttaaacad tatgatgtgt gaaaactggg 240  
 taatatttat aggtcacttt gttttactgt cttaagttta tactcttata gacaacatgg 300  
 ccgtgaactt tatgctgtaa ataatcagag gggaataaac tgttgagtca aaacagccat 360  
 cttccttgtg accaaacatt taaaaatatt ctggctgggc acagtggctc acgcctgtaa 420  
 tcccagcact ttgggaggcc gaggtgggca gatcacctga ggttggg 467

<210> 34  
 <211> 460  
 <212> DNA  
 <213> Homo sapiens

<400> 34  
 tgacaacaga cttctgaaat agaccacaa gaggcagttc catttcattt gtgccagaat 60  
 gctttaggat gtacagttaa ggattgaaag tttacaggaa aaaaaattag gccgttcctt 120  
 caaagcaaat gtcttcctgg attattcaaa atgatgtatg ttgaagcctt tgtaaatgtt 180  
 cagatgctgt gcaaagttaa ttattttaaa cattatgatg tgtgaaaact ggttaatat 240  
 tatagatcac tttgttttac tgtcttaagt ttatactctt atagacaaca tggccgtgaa 300  
 ctttatgctg taaataatca gaggggaata aactgttgag tcaaaacagc catcttcctt 360  
 gtgaccaaac atttaaaaat attctggctg ggcacagtgg ctacgcctg taatcccagc 420  
 actttgggag gccgaggttg gcagatcacc tgaggttggg 460

<210> 35  
 <211> 462  
 <212> DNA  
 <213> Homo sapiens

<400> 35  
 tctgacaaca gacttctgaa atagaccac aagaggcagt tccatttcat ttgtgccaga 60  
 atgctttagg atgtacagtt atggattgaa agtttacagg aaaaaaatt aggccttcc 120  
 ttcaaagcaa atgtcttctt ggattattca aaatgatgta tgttgagcc tttgtaaatt 180  
 gtcagatgct gtgcaaatgt tattatttta aacattatga tgtgtgaaaa ctgggttaata 240  
 tttatagatc actttgtttt actgtcttaa gtttatactc ttatagacaa catggccgtg 300  
 aactttatgc tgtaataaat cagaggggaa taaactgttg agtcaaaaca gccatcttcc 360  
 ttgtgaccaa acatttaaaa atattctggc tgggcacagt ggctcacgcc tgtaatccca 420  
 gcactttggg aggcagaggt gggcagatca cctgaggttg gg 462

<210> 36  
 <211> 463  
 <212> DNA  
 <213> Homo sapiens

<400> 36  
 ctctgacaac agacttctga aatagacca caagaggcag ttccatttca tttgtgccag 60

- 13 -

```

aatgctttag gatgtacagt tatggattga aagtttacag gaaaaaaaaat taggccgttc 120
cttcaaagca aatgtcttcc tggattattc aaaatgatgt atgttgaagc ctttgtaaata 180
tgtcagatgc tgtgcaaata ttattatttt aaacattatg atgtgtgaaa actgggttaat 240
atttatagrt cactttgttt tactgtctta agtttatact cttatagaca acatggccgt 300
gaactttatg ctgtaaataa tcagagggga ataaactgtt gagtcaaaac agccatcttc 360
cttgtgacca aacattttaa aatattctgg ctgggcacag tggctcacgc ctgtaatccc 420
agcactttgg gaggccgagg tgggcagatc acctgaggtt ggg 463

```

&lt;210&gt; 37

&lt;211&gt; 17

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 37

ggtggctggg ctcttct

17

&lt;210&gt; 38

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 38

ctcgcttcct cagtacctat gatg

24

&lt;210&gt; 39

&lt;211&gt; 21

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 39

ctggctgagt gccagacatc t

21

&lt;210&gt; 40

&lt;211&gt; 17

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 40

ggcgggatgg agtggaa

17

&lt;210&gt; 41

&lt;211&gt; 21

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 41

ccacctcaag ctctggtgat c

21

&lt;210&gt; 42

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

- 14 -

<400> 42  
gttgactctt ttggcctttt cag

23

<210> 43  
<211> 23  
<212> DNA  
<213> Homo sapiens

<400> 43  
ccttaccaga cttccaggat ggt

23

<210> 44  
<211> 25  
<212> DNA  
<213> Homo sapiens

<400> 44  
tgtccaataa ctgcatcacc tacct

25

<210> 45  
<211> 13  
<212> DNA  
<213> Homo sapiens

<400> 45  
catggctgga ccc

13

<210> 46  
<211> 13  
<212> DNA  
<213> Homo sapiens

<400> 46  
catggctgga tcc

13

<210> 47  
<211> 13  
<212> DNA  
<213> Homo sapiens

<400> 47  
tgctccggcg cca

13

<210> 48  
<211> 14  
<212> DNA  
<213> Homo sapiens

<400> 48  
ctgctctggc gccca

14

<210> 49  
<211> 16

- 15 -

<212> DNA  
<213> Homo sapiens

<400> 49  
ctctgttgcc ccagaa 16

<210> 50  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 50  
ctctgttgcg ccaga 15

<210> 51  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 51  
ctttcaaggg cctgc 15

<210> 52  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 52  
cctttcaagg ggcct 15

<210> 53  
<211> 16  
<212> DNA  
<213> Homo sapiens

<400> 53  
aagactcgag tgcct 16

<210> 54  
<211> 16  
<212> DNA  
<213> Homo sapiens

<400> 54  
agactcaagt gtcctc 16

<210> 55  
<211> 533  
<212> DNA  
<213> Homo sapiens

<400> 55  
ttcgtctcag tttgtttgtg agcaggctgt gagtttgggc cccagaggct ggggtgacatg 60



- 16 -

```

tggtggcagc ctcttcaaaa tgagccctgt cctgcctaag gctgaacttg ttttctggga 120
acaccatagg tcacctttat tctggcagag gaggagcat cagtgccttc caggatagac 180
ttttcccaag cctacttttg ccattgactt cttcccaaga ttcaatccca ggatgtacaa 240
ggacagcccc tctcccatag tatgggactg gcctctgctg atcctcccag gcttcogtgt 300
gggtcagtg ggcctatgga tgtgcttggt aactgagtgc cttttggtgg agaggcccg 360
cctctcaciaa aagacccttt accactgctc tgatgaagag gactacacag aacacataat 420
tcaggaagca gctttcccca tgtctcgact catccatcca ggccattccc cgtctctggt 480
tcctcccttc ctctgggact cctgcacacg ctccttcttc tgaggctgaa att 533

```

&lt;210&gt; 56

&lt;211&gt; 497

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 56

```

gggccccaga ggctgggtga catgtgttgg cagcctcttc aaaatgagcc ctgtcctgcc 60
taaggctgaa cttgttttct gggaacacca taggtcacct ttattctggc agaggaggga 120
gcatcagtg cctccaggat agacttttcc caagcctact ttggccattg acttcttccc 180
aagattcaat cccaggatgt acaaggacag cccctcctcc atagtatggg actggcctct 240
gctgatcctc ccaggcttcc gtgtgggtca gtggggccca tggatgtgct tgttaactga 300
gtgccttttg gtggagaggc cgggcctctc acaaaaagacc ccttaccact gctctgatga 360
agaggagtac acagaacaca taattcagga agcagctttc cccatgtctc gactcatcca 420
tccaggccat tccccgtctc tggttcctcc cctcctcctg gactcctgca cacgctcctt 480
cctctgaggc tgaaatt 497

```

&lt;210&gt; 57

&lt;211&gt; 497

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 57

```

gggccccaga ggctgggtga catgtgttgg aagcctcttc aaaatgagcc ctgtcctgcc 60
taaggctgaa cttgttttct gggaacacca taggtcacct ttattctggc agaggaggga 120
gcatcagtg cctccaggat agacttttcc caagcctact ttggccattg acttcttccc 180
aagattcaat cccaggatgt acaaggacag cccctcctcc atagtatggg actggcctct 240
gctgatcctc ccaggcttcc gtgtgggtca gtggggccca tggatgtgct tgttaactga 300
gtgccttttg gtggagaggc cgggcctctc acaaaaagacc ccttaccact gctctgatga 360
agaggagtac acagaacaca taattcagga agcagctttc cccatgtctc gactcatcca 420
tccaggccat tccccgtctc tggttcctcc cctcctcctg gactcctgca cacgctcctt 480
cctctgaggc tgaaatt 497

```

&lt;210&gt; 58

&lt;211&gt; 497

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 58

```

gggccccaga ggctgggtga catgtgttgg cagcctcttc aaaatgagcc ctgtcctgcc 60
taaggctgaa cttgttttct gggaacacca taggtcacct ttattctggc agaggaggga 120
gcatcagtg cctccaggat agacttttcc caagcctact ttggccattg acttcttccc 180
aagattcaat cccaggatgt acaaggacag cccctcctcc atagtatggg actggcctct 240
gctgatcctc ccaggcttcc gtgtgggtca gtggggccca tggatgtgct tgttaactga 300
gtgccttttg gtggagaggc cgggcctctc acaaaaagacc ccttaccact gctctgatga 360

```

- 17 -

agaggagtag acagaacaca taattcagga agcagctttc cccatgtctc gactcatcca 420  
 tccaggccat tccccgtctc tgggtccctc cctcctcctg gactcctgca cacgctcctt 480  
 cctctgaggg tgaaatt 497

&lt;210&gt; 59

&lt;211&gt; 483

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 59

gggtgacatg tgttggcagc ctcttcaaaa tgagccctgt cctgcctaag gctgaacttg 60  
 ttttctggga acaccatagg tcacctttat tctggcagag gagggagcat cagtgccttc 120  
 caggatagac ttttcccaag cctacttttg ccattgactt cttcccaaga ttcaatccca 180  
 ggatgtacaa ggacagcccc tctccatag tatgggactg gcctctgctg atcctcccag 240  
 gcttccgtgt gggtcagtgg ggcccatgga tgtgcttgtt aactgagtgc cttttggtgg 300  
 agaggcccg cctctcaca aagaccctt accactgtct tgatgaagag gactacacag 360  
 aacacmtaat tcaggaagca gctttcccca tgtctcgact catccatcca ggccattccc 420  
 cgtctctggg tctccctc ctcttggtgact cctgcacacg ctccttctc tgaggctgaa 480  
 att 483

&lt;210&gt; 60

&lt;211&gt; 500

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 60

tttgggcccc agaggctggg tgacatgtgt tggcagcctc ttcaaatga gccctgtcct 60  
 gcctaaggct gaacttgttt tctgggaaca ccatagggtca cctttattct ggcagaggag 120  
 ggagcatcag tgccctccag gatagacttt tcccaagcct acttttgcca ttgacttctt 180  
 cccaagattc aatcccagga tgtacaagga cagccctctc tccatagtat gggactggcc 240  
 tctgtgatc ctcccaggct tccgtgtggg tcagtggggc ccattggatgt gcttggttaac 300  
 tgagtgcctt ttggtggaga ggcccgccct ctcaaaaag accccttmcc actgctctga 360  
 tgaagaggag tacacagaac acataattca ggaagcagct ttcccatagt ctgactcat 420  
 ccatccaggc cattccccgt ctctggttcc tccctctc ctggactcct gcacacgctc 480  
 cttcctctga ggctgaaatt 500

&lt;210&gt; 61

&lt;211&gt; 499

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 61

ttgggcccc gaaggctggg gacatgtgtt ggcagcctct tcaaatgag ccctgtcctg 60  
 cctaaggctg aacttgtttt ctgggaacac cataggtcac ctttattctg gcagaggagg 120  
 gagcatcagt gccctccagg atagactttt cccaagccta cttttgcat tgacttcttc 180  
 ccaagattca atcccaggat gtacaaggac agccctcct ccattagtat ggactggcct 240  
 ctgtgatcc tcccaggctt ccgtgtggg cagtggggc catggatgtg cttgttaact 300  
 gagtgcctt tgggtggagag gcccgccctc tcacaaaaga ccccttmcca ctgctctgat 360  
 gaagaggagt acacagaaca cataattcag gaagcagctt tcccatgtc tgcactcatc 420  
 catccaggcc attccccgtc tctggttctt cctcctctc tggactcctg cacacgctcc 480  
 ttctctgag gctgaaatt 499

&lt;210&gt; 62

- 18 -

<211> 498  
 <212> DNA  
 <213> Homo sapiens

<400> 62  
 tgggccccag aggctgggtg acatgtgttg gcagcctctt caaaatgagc cctgtcctgc 60  
 ctaaggctga acttgttttc tgggaacacc ataggtcacc tttattctgg cagaggaggg 120  
 agcatcagtg ccctccagga tagacttttc ccaagcctac ttttgccatt gacttcttcc 180  
 caagattcaa tcccaggatg tacaaggaca gcccctcctc catagtatgg gactgggctc 240  
 tgctgatcct cccaggcttc cgtgtgggtc agtggggccc atggatgtgc ttgttaactg 300  
 agtgcccttt ggtggagagg cccggcctct cacaaaagac cccttmccac tgctctgatg 360  
 aagaggagta cacagaacac ataattcagg aagcagcttt ccccatgtct cgactcatcc 420  
 atccaggcca ttccccgtct ctggttcctc ccctcctcct ggactcctgc acacgctcct 480  
 tcctctgagg ctgaaatt 498

<210> 63  
 <211> 498  
 <212> DNA  
 <213> Homo sapiens

<400> 63  
 tgggccccag aggctgggtg acatgtgttg gcagcctctt caaaatgagc cctgtcctgc 60  
 ctaaggctga acttgttttc tgggaacacc ataggtcacc tttattctgg cagaggaggg 120  
 agcatcagtg ccctccagga tagacttttc ccaagcctac ttttgccatt gacttcttcc 180  
 caagattcaa tcccaggatg tacaaggaca gcccctcctc catagtatgg gactggcctc 240  
 tgctgatcct cccaggcttc cgtgtgggtc agtggggccc atggatgtgc ttgttaactg 300  
 agtgcccttt ggtggagagg cccggcctct cacaaaagac cccttmccac tgctctgatg 360  
 aagaggagta cacagaacac ataattcagg aagcagcttt ccccatgtct cgactcatcc 420  
 atccaggcca ttccccgtct ctggttcctc ccctcctcct ggactcctgc acacgctcct 480  
 tcctctgagg ctgaaatt 498

<210> 64  
 <211> 498  
 <212> DNA  
 <213> Homo sapiens

<400> 64  
 tgggccccag aggctgggtg acatgtgttg gcagcctctt caaaatgagc cctgtcctgc 60  
 ctaaggctga acttgttttc tgggaacacc ataggtcacc tttattctgg cagaggaggg 120  
 agcatcagtg ccctccagga tagacttttc ccaagcctac ttttgccatt gacttcttcc 180  
 caagattcaa tcccaggatg tacaaggaca gcccctcctc catagtatgg gactggcctc 240  
 tgctgatcct cccaggcttc cgtgtgggtc agtggggccc atggatgtgc ttgttaactg 300  
 agtgcccttt ggtggagagg cccggcctct cacaaaagac cccttaccac tgctctgatg 360  
 aagaggagta cacagaacac ataattcagg aagcagcttt ccccatgtct cgactcatcc 420  
 atccaggcca ttccccgtct ctggttcctc ccctcctcct ggactcctgc acacgctcct 480  
 tcctctgagg ctgaaatt 498

<210> 65  
 <211> 503  
 <212> DNA  
 <213> Homo sapiens

<400> 65

- 19 -

```

gagtttgggc cccagaggct ggggtgacatg tgttggcagc ctcttcaaaa tgagccctgt 60
cctgcctaag gctgaacttg ttttctggga acaccatagg tcacctttat tctggcagag 120
gagggagcat cagtgccttc caggatagac ttttcccaag cctacttttg ccattgactt 180
cttcccaaga ttcaatccca ggatgtacaa ggacagcccc tcctccatag tatgggactg 240
gcctctgctg atcctcccag gcttccgtgt gggtcagtgg ggcccatgga tgtgcttggt 300
aactgagtgc cttttggtgg agaggcccg cctctcacia aagaccctt cccactgctc 360
tgatgaagag gactacacag aacacataat tcaggaagca gctttcccca tgtctcgact 420
catccatcca ggccattccc cgtctctggt tcctcccctc ctcttggtgact cctgcacacg 480
ctccttcttc tgaggctgaa att 503

```

&lt;210&gt; 66

&lt;211&gt; 453

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 66

```

tgagccctgt cctgcctaag gctgaacttg ttttctggga acaccatagg tcacctttat 60
tctggcagag gagggagcat cagtgccttc caggatagac ttttcccaag cctacttttg 120
ccattgactt cttcccaaga ttcaatccca ggatgtacaa ggacagcccc tcctccatag 180
tatgggactg gcctctgctg atcctcccag gcttccgtgt gggtcagtgg ggcccatgga 240
tgtgcttggt aactgagtgc cttttggtgg agaggcccg cctctcacia aagaccctt 300
cccactgctc tgatgaagag gactacacag aacacataat tcaggaagca gctttcccca 360
tgtctcgact catccatcca ggccattccc cgtctctggt tcctcccctc ctcttggtgact 420
cctgcacacg ctccttcttc tgaggctgaa att 453

```

&lt;210&gt; 67

&lt;211&gt; 500

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 67

```

tttgggcccc agaggctggg tgacatgtgt tggcagcctc ttcaaaatga gccctgtcct 60
gcctaaggct gaacttggtt tctgggaaca ccatagggtca cctttattct ggcagaggag 120
ggagcatcag tgccctccag gatagacttt tcccaagcct acttttgcca ttgacttctt 180
cccaagattc aatcccagga tgtacaagga cagcccctcc tccatagtat gggactggcc 240
tctgctgata ctcccaggct tccgtgtggg tcagtggggc ccatggatgt gcttggttaac 300
tgagtgcctt ttggtggaga ggcccggcct ctcaaaaaag accccttmcc actgctctga 360
tgaagaggag tacacagaac acataattca ggaagcagct ttcccatgt ctcgactcat 420
ccatccaggc cattccccgt ctctggttcc tcccctcctc ctggactcct gcacacgctc 480
cttctcttga gggtgaaatt 500

```

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☒ **BLACK BORDERS**

☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**

☐ **FADED TEXT OR DRAWING**

☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**

☐ **SKEWED/SLANTED IMAGES**

☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**

☐ **GRAY SCALE DOCUMENTS**

☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**

☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**

☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**